ABSTRACT

Title of Dissertation: NEXT-GENERATION MASS SPECTROMETRY WITH MULTI-OMICS FOR DISCOVERIES IN CELL AND NEURODEVELOPMENTAL BIOLOGY

Jie Li, Doctor of Philosophy, 2022

Dissertation directed by: Prof. Peter Nemes, Department of Chemistry and Biochemistry

Understanding tissue formation advances our understanding of the causes of disease and the obtained knowledge can be potentially applied to develop personalized interventions. However, to explore the underlying mechanisms that govern tissue formation, there is a high and unmet need to develop new technologies to characterize different types of biomolecules from early-stage embryonic precursor cells and their descendent cells during development. This dissertation discusses new technological advancements to facilitate multi-omic (proteomic and metabolomic) analysis to explore cell-to-cell differences and uncover mechanisms underlying tissue formation. The work presented herein illustrates the development of *in vivo* microsampling and single-cell mass spectrometry (MS) to uncover cell heterogeneity among embryonic cells. Additionally, this dissertation work studies the biological role of metabolites in cell fate determination by exploring the mechanisms underlying metabolite-induced cell fate change. Moreover, this work introduces a novel technique called MagCar developed to track and isolate tissue-specific cells at later stages, which enables studying temporal molecular changes to gain new information about tissue formation.

Chapter 1 overviews key aspects of the technologies that are integrated into this dissertation, including high resolution-mass spectrometry (HRMS), liquid chromatography (LC), capillary electrophoresis (CE), multi-omics, research model, research significance, and research motivation.

Chapter 2 extends my analysis tools to study systems and functional biology.

Chapter 3 presents the development of the first dual proteo-metabolomic single-cell HRMS for *in vivo* analysis in live frog embryos.

Chapter 4 describes FACS-based tissue-specific cell isolation along with LC-HRMS-based metabolomics and proteomics to uncover molecular reorganization underlying metabolite-induced cell fate change.

Chapter 5 introduces a novel sampling technique, called MagCar, to isolate tissue-specific cell lineages from live frog embryos for temporal studies.

Chapter 6 summarizes the research accomplishments in this dissertation and prospects for future directions to further advance research in understanding tissue formation.

NEXT-GENERATION MASS SPECTROMETRY WITH MULTI-OMICS FOR DISCOVERIES IN CELL AND NEURODEVELOPMENTAL BIOLOGY

by

Jie Li

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Associate Professor Peter Nemes, Chair
Associate Professor Kan Cao, Dean’s Representative
Professor Neil V. Blough
Distinguished University Professor Catherine Fenselau
Assistant Professor Nishanth E. Sunny
Professor Lai-Xi Wang

Dedication

To my parents, family, and girlfriend Grace for unwavering support and encouragement over the years.

Acknowledgements

It finally comes to the point of graduation. When I recall the first day I started my Ph.D. study, I would never believe I can come to where I am now. The Ph.D. journey would not be such smooth if I didn’t get mentorships, help, support, encouragement, and suggestions from so many people that I cannot name all of them.

I feel very lucky to join Nemes Research Group. I would like to express my sincere gratitude and thanks to my advisor Prof. Peter Nemes for allowing me to conduct dissertation research in his lab. It’s a lot of fun to get involved in exciting, interdisciplinary, and cutting-edge research projects to hone my various skills. Moreover, I appreciate the opportunities offered by Prof. Nemes to develop my scientific communication skills by attending local and national conferences to present my research. I am grateful for his mentorship, patience, and support which finally prepare me to be a scientist with the skillset that will be a solid foundation for achieving my long-term career goals.

I am grateful to Prof. Catherine Fenselau, Prof. Neil V. Blough, Prof. Nishanth E Sunny, Prof. Lai-Xi Wang, and Prof. Kan Cao for your time and input while serving on my dissertation committee. I would like to thank Prof. Sally Moody from GWU and her lab members for allowing me to work in her lab to receive the initial training on embryology. I want to thank Migo Gui, Faith Isaac Johnson, Diane Canter, Melanna Mills, Prof. Neil Blough, and Prof. Efrain E. Rodriguez from the Graduate Study Office of the UMD Department of Chemistry and Biochemistry for help with the administration and logistics and for navigating me towards graduation.

I would like to express my thanks to Dr. Camille Lombard-Banek, Dr. Erika Portero, Dr. Rosemary Onjiko, Dr. Sam Choi, Dr. Aparna Baxi, and Vi Quach who trained me on different topics and helped me get started when I first joined the lab. I want to thank Dr. Kaitlyn Stepler who collaborated with me on various projects presented in this dissertation to move forward with my research. I also want to thank all other Nemes Research Group members including Dr. Kellen Delaney, Leena Pade, Dr. Chathuri Mohottige, Dashuang Jia, Ryan O'Neal, Bowen Shen, John Mendis, Steven Woolford, Ailing Li, Fei Zhou, Yayra Tuani, Tarikul Islam, Chase Singer, and Rohith Mathew. The dissertation research will not go as smoothly as this without input from all the group members which include but are not limited to providing feedback, critiques, and suggestions in group meetings and talk rehearsals and reviewing my abstracts, proposals, and manuscripts. I am also grateful to Ken Class from UMD Flow Cytometry Core Facility for help with FACS experiments and Amy Beaven from UMD Imaging Core Facility for help with confocal microscopy imaging. I would like to thank Dr. Yue Li from Mass Spectrometry Facility at UMD for helping with access to the maXis II mass spectrometer used in this dissertation work.

The research described in this dissertation was funded by National Science Foundation (NSF), National Institutes of Health (NIH), Cosmos Club Foundation, and Arnold and Mabel Beckman Foundation. I would like to thank all the funding agencies and sources for making this research financially feasible. I also want to thank the Washington-Baltimore Mass Spectrometry Discussion Group (WBMSDG) and the Washington Chromatography Discussion Group (WCDG) for the awards that supported my travels to conferences. Moreover, I appreciate the UMD Graduate School for the Ann G. Wylie Dissertation Fellowship, which supported me through the final semester of my dissertation research to write this dissertation.

Table of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Abbreviations

CE	Capillary electrophoresis
CID	Collision-induced ionization
D1	Dorsal-animal cell
D11	Dorsal-animal midline cell
DDA	Data-dependent acquisition
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FACS	Fluorescence-activated cell sorting
FC	Fold change
FMP	Fluorescent magnetic particle
HCD	Higher-energy collisional dissociation
HILIC	Hydrophilic interaction chromatography
HRMS	High-resolution mass spectrometry
IMS	Ion mobility separation
LC	Liquid chromatography
m/z	Mass-to-charge ratio
MACS	Magnetic-activated cell sorting
MagCar	Cell magnetic cargo for cell sorting
MT	Migration time
OT	Orbitrap
ppm	Parts per million
Q	Quadrupole
RP	Reverse phase
RSD	Relative standard deviation
RT	Retention time
S/N	Signal-to-noise ratio
TMT	Tandem mass tags
TOF	Time-of-flight
V1	Ventral-animal cell
V11	Ventral-animal midline cell

Chapter 1: Introduction

To unravel the mechanisms underlying various biological activities such as cancer diseases and neuroscience, research focuses on the role of functionally and biologically important biomolecules. Our understanding of the cell, the functional and structural building block of life, depends on the availability of technologies that can measure all the different types of biomolecules that the cell makes. These biomolecules, including upstream RNAs and DNAs and downstream proteins and metabolites, have distinct roles and interact through complex molecular networks. For example, RNAs participate in the coding, regulation, and expression of genetic information carriers—DNAs. Proteins, translated from RNAs, have diverse roles such as structural elements of cells, enzymes to catalyze biochemical reactions, transcription factors, and receptors of signaling molecules to respond to intracellular and extracellular changes. Metabolites, which act downstream to proteins and hence are indicators of cell phenotypic state, were found not only to be effective biomarkers but also active drivers and modulators of biological processes.[1] Studying these different biomolecules at the level of small-to-large samples has the potential to provide insight into the mechanisms involved in cell and developmental processes underlying health and disease states.

1.1 HRMS for metabolomics and proteomics

Several techniques can detect downstream proteins and metabolites in cells, tissues, and organisms but not equally as well. For example, nuclear magnetic resonance (NMR) is widely used for metabolomic and proteomic analysis due to its non-destructive characteristics[2] and structure elucidation capability[3]. However, the application of NMR for 'omics is limited, because NMR has a relatively lower sensitivity compared with high-resolution mass spectrometry

(HRMS)[2] and requires a larger starting sample amount. HRMS is the benchmark for detecting and identifying biomolecules from small-to-large amounts of materials as it is label-free and can provide structural information to support confident identification.[4-5] HRMS offers high sensitivity, enabling the detection of low-abundance biomolecules in single cells, such as embryonic cells, neurons, and tissues that my research focused on. For example, the quadruple time-of-flight (Q-TOF) HRMS allowed me to detect neurotransmitter metabolite acetylcholine with a limit of detection (LOD) < 50 nM,[6] providing enough sensitivity to detect many other endogenous metabolites with concentrations ranging from ~µM–mM[7] in biological systems. My research also advances proteomics using Orbitrap HRMS, as demonstrated by the ultrasensitive detection of brain angiotensin peptides with a lower LOD of ~260 zmol.[8] The performance metrics of Q-TOF and Orbitrap HRMS were summarized according to a recent review paper, which shows that Q-TOF HRMS has a wider dynamic range and mass range while Orbitrap HRMS has a higher resolution.[9]

Moreover, HRMS provides high confidence in compound identification[10] due to the high resolution[9] to distinguish analytes with similar mass-to-charge ratio (*m/z*) and the capability of the MS/MS fragmentation. The availability of multiple MS/MS databases[11-12] with an increasing number of reference metabolites as well as open-source metabolomic data processing software including XCMS online[13] and MZmine[14] expedites the application of mass spectrometry in metabolomics. Likewise, software including MaxQuant[15-16] and Proteome Discoverer[17] was developed to help with protein identification and quantification from proteomic MS data.

Development in instrumentation advanced the application of MS to even smaller amounts of biomolecules and materials. For example, the recent advent of ion mobility separation (IMS) techniques decreases the complexity of samples by providing post-ionization separation of ions based on their ion mobility difference.[18-19] As a result, the lower LOD of TOF HRMS was further decreased to 30 zmol[20] by incorporating IMS to remove chemical noise[18] compared with a lower LOD of 260 zmol[8] without IMS. Moreover, the IMS technique increases the identification confidence of biomolecules by measuring an intrinsic property called collision cross section (CCS)[21] along with accurate mass and MS/MS fragmentation pattern. IMS coupling to MS has been applied in isomer and isobaric ions separation.[18, 21]

HRMS allowed me to quantify proteomic and metabolomic states. For metabolites, under-the-curve peak areas were used as a quantitative proxy to compare the relative metabolite levels of samples under different conditions. For proteins, the tandem mass tags (TMT) technique provided multiplexity capability for bottom-up proteomics, enabling the acquisition of high-quality MS data for ‘omic analysis and improving the depth of MS-based identification[22], data quality, and intra- and inter-laboratory reproducibility.[23-24] Moreover, TMT enables the quantification of proteins from multiple samples simultaneously, which shortens measurement time to achieve several folds higher throughput.[23-24] Furthermore, research has found that TMT can increase the number of quantified proteins and improve the reliability of relative quantification compared with label-free quantification.[24] Application of the TMT labeling strategy has led to the identification of more disease biomarkers in early myocardial injury,[24] for example. TMT also enhanced protein identification in neural ectoderm in *Xenopus laevis* (*X. laevis)* embryos to help understand the development of the central nervous system.[22]

Metabolomics and proteomics based on MS have advanced research in cell and developmental biology and made it possible to delve into complicated biological questions involved in biological systems at different levels including organisms[7], microorganisms[25], tissues[22], single cells[26-27], and subcellular organelles[28]. For example, metabolomics based on MS found metabolic remodeling during the cleavage and mid-blastula embryonic stages at the whole-embryo level for *Xenopus* frogs.[7] In my research, MS methods were developed to detect and identify metabolites and proteins from isolated tissue-specific cell populations, which uncovered molecular reorganization during metabolite-induced cell fate change actions. I also developed single-cell MS tools for both metabolomics and proteomics at the single-cell level, which I applied to reveal cell-to-cell differences along the dorsal-ventral axis.[26] Other examples are the application of MS in the analysis of proteins and metabolites in single cells to explore cell heterogeneity.[6, 29-30] A typical metabolomics[31] and proteomics workflow[32] was summarized in **Figure 1.1**. The proteomics community recently recommended guidelines in experimental practices and data reporting for proteomics[33]. These different MS-based ‘omics techniques discovered small molecules that caused embryonic cell fate change[6], studied causes of cancer disease[34], and facilitated disease biomarker discovery[24, 35].

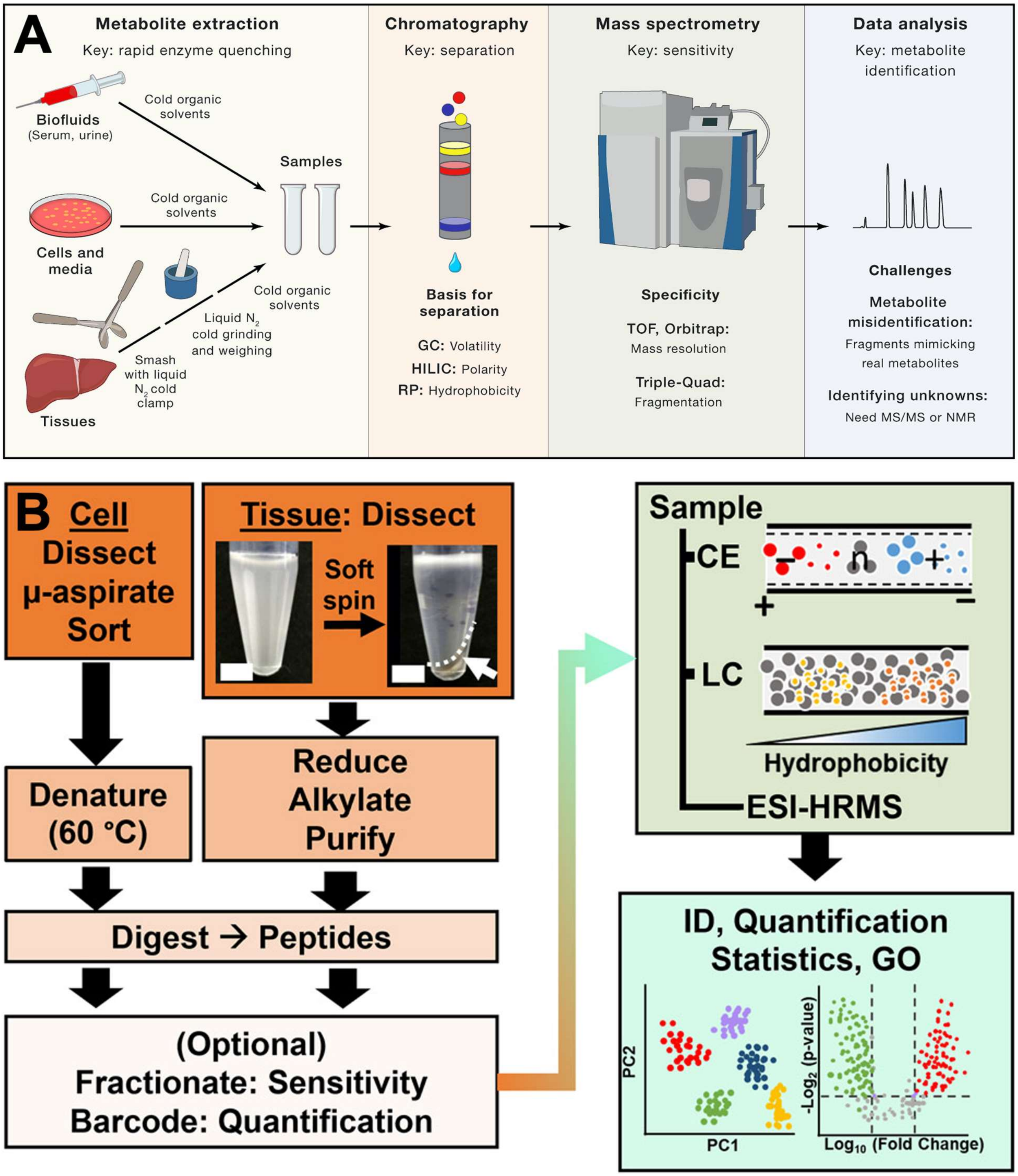


Figure 1.1 General workflows for MS-based metabolomics and proteomics. (**A**) A typical MS-based metabolomics workflow for various types of biological samples from ref. 31. Reprinted from *Cell*, 173, Rabinowitz, J. D *et al*, Metabolomics and Isotope Tracing, 822–837, Copyright (2018), with permission from Elsevier. (**B**) A typical MS-based bottom-up proteomics workflow for various types of biological samples. Reproduced from ref. 32 with permission from JoVE.

1.2 Separation techniques compatible with MS: LC and CE

To help biomolecular measurements, MS is commonly coupled with separation techniques such as capillary electrophoresis (CE) and liquid chromatography (LC) to detect proteins and metabolites in complex sample matrices. Separation decreases the complexity of samples to mitigate ion suppression during ionization and spectral interferences during MS detection. Separation also introduces separation time as an orthogonal parameter for compound identification.[36] When using MS hyphenated to separation techniques for identification, a confident identification requires the molecular feature under investigation to match the accurate mass, isotopic pattern, MS/MS fragmentation, and separation time as recommended by the metabolomics community.[12] A typical LC-MS workflow for metabolite identification was summarized in **Figure 1.2**.[31]

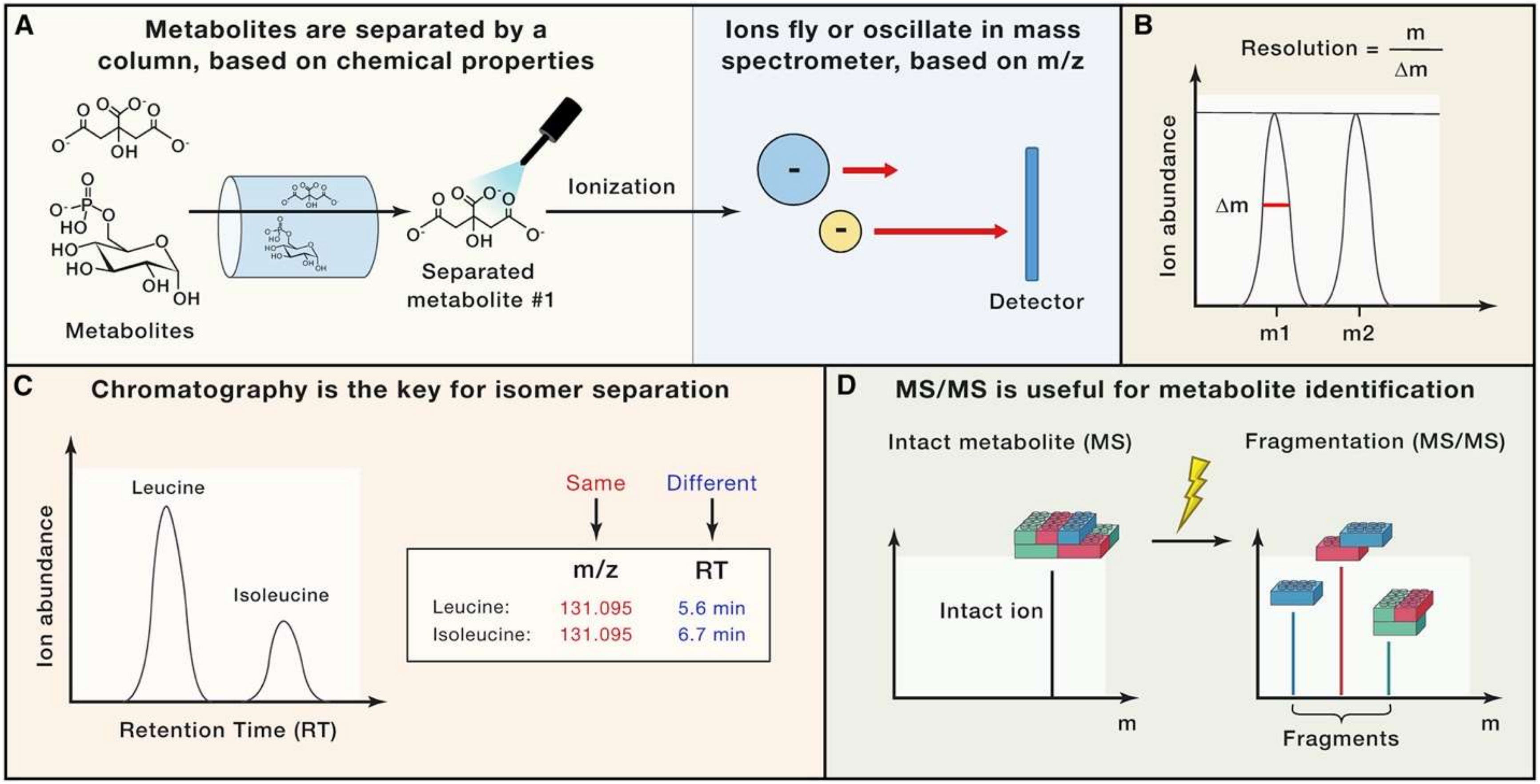


Figure 1.2 A typical LC-MS workflow for detecting and identifying metabolites using accurate mass, retention time, and MS/MS comparison from ref. 31. Reprinted from *Cell*, 173, Rabinowitz, J. D *et al*, Metabolomics and Isotope Tracing, 822–837, Copyright (2018), with permission from Elsevier.

LC coupled to MS is the gold standard separation technology for metabolomic and proteomic analysis because it allows for the simultaneous detection of large numbers of analytes from small samples with high throughput.[12, 37-38] Separation of analytes in LC is based on the interaction between the analytes and the stationary phase of the column. Depending on the polarity of the LC column's stationary phase, LC may be grouped into the area of reversed-phase LC (RPLC) and hydrophilic-interaction chromatography (HILIC). RPLC is suitable for separating less polar analytes because the nonpolar stationary phase allows for better retention of hydrophobic analytes such as peptides and proteins. In contrast, polar analytes have better separation using HILIC.[37] Recently, developments of new types of LC columns have put the application of LC forward. For example, micropillar array columns (μPAC), based on microlithographic etching to generate highly ordered and reproducible pillars in a chip, were recently developed to minimize backpressure, peak dispersion, retention time variation, and column-to-column difference in LC separation.[39-40] Unlike the packed HPLC column, the separation efficiency of which is limited by the eddy dispersion (multi-path term, A, in **van Deemter equation**), μPAC as an open tubular LC column is free of eddy dispersion contribution. The plate number of a 200-cm μPAC column is 300,000 which is around one order of magnitude compared with that of packed HPLC columns (≤20,000).[40] Due to the above-mentioned advantages compared to packed LC column, my research adopted μPAC to enable the identification of more proteins. μPAC has been applied in various areas including proteomics, pharmaceutical research, and lipidomics.[21, 39-40]

$$H = A + \frac{B}{\mu} + C\mu \qquad \textbf{(van Deemter equation)}$$

Where H is plate height, A is the multi-path term, $(\frac{B}{\mu})$ is the longitudinal diffusion term, and $(C\mu)$ is the mass-transfer term.

For larger amounts of starting materials, LC offers advantages in throughput, post-acquisition data processing, and scalability. In my research to measure tissue-specific cell populations, I used LC because the LC autosampler enables the temperature control of samples and automatic sample injection, thus allowing queueing samples for continuous measurement to shorten experiment time. Moreover, the reproducible sample injection volume and analyte separation reduce the complexity of data processing after the data acquisition, which is required in CE including alignment of separation time[41-42]. Additionally, LC has the advantage of being able to tailor to various physicochemical properties of analytes to drive separation. The metabolites extracted from *X. laevis* embryos were previously revealed to be majorly polar metabolites including amino acids and neurotransmitters.[43] Therefore, I chose HILIC to achieve better separation of analytes from the metabolite extract of collected cell populations from *X. laevis* embryos. In contrast, the peptides resulting from trypsin digestion are less polar and, therefore, RPLC is commonly used for bottom-up proteomics[44-45] and was chosen for proteomics in my research.

CE, a separation technique that requires less sample compared with LC, supported my research in single-cell analysis with compatibility with limited amounts of materials available in single cells. CE separates analytes based on their electrophoretic velocity differences, determined by the size (hydrated radius, r) and charge state (z) of analytes, under the applied high voltage. Generally, the greater the z/r ratio, the greater the electrophoretic mobility. Analyte movement in the capillary is also under the effect of bulk solution movement caused by electro-osmotic flow (EOF). For example, the fused silica capillary forms a surface-bound Si-O$^-$ under a background electrolyte (BGE) solution with a pH >2. The negative surface charge results in the adsorption of cations from BGE to form a tight layer of ions on the surface of the inner capillary, which is rich in anions. The negative charge in turn results in the formation of a diffusion layer that is rich in cations, thus, a double layer is formed. Under the applied high voltage, cations in the diffusion layer move to the cathode, which brings the bulk solution to move along toward the cathode, and thus, the EOF is formed. Under the combined effect of the electrophoretic mobility and EOF, analytes are separated and eluted out of the CE capillary. Since EOF depends on the pH of BGE, EOF can be deliberately manipulated by changing the pH of BGE for specific applications of CE. **Figure 1.3** illustrates the separation of charged analytes with different hydrated radii and charges.

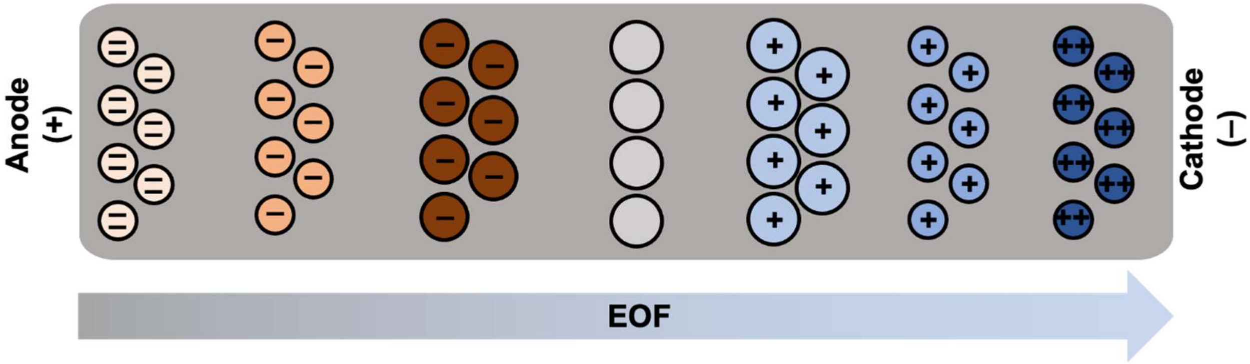


Figure 1.3 Illustration of the CE separation of ions with different hydrated radii and charge under the combined effect of electrophoretic velocity and EOF. Blue circles are positively charged ions and brown circles are negatively charged ions. A larger circle size means a larger hydrated radius. Smaller ions with higher positive charges are eluted out of the cathode first under the applied high voltage.

CE has high separation efficiency which comes from the large theoretical plate number (N). This is because the peak broadening (H, plate height) in CE is minimized as the peak broadening is free of contribution from the multi-path term (A) and mass-transfer term ($C\mu$), and H is solely contributed by the longitudinal diffusion term ($\frac{B}{\mu}$) in **van Deemter equation**.[46] The high separation efficiency of CE results in narrow peaks in separation, which benefits the signal-to-noise ratio (S/N) for ultrasensitive measurement. For instance, our custom-built CE-ESI-MS system achieved a sensitivity of ~60 amol for metabolites and ~700 zmol for peptides.[26]

Moreover, CE is compatible with minute sample volumes down to the nanoliters available in a single cell,[47] which makes CE-MS widely used in the single-cell analysis by integrating with MS.[47-48] For example, single embryonic cells from the South African frog *X. laevis* embryos have a relatively large cell size (~250 µm in diameter at 16-cell stage) but limited total cellular contents (~100 nL), which is compatible with CE. In contrast, the available cellular contents in single cells are not sufficient for the typical micro-flow LC, which requires the starting amounts of samples to be at µL scale—10 times the available cellular contents in single cells. Therefore, in my research to understand cell-to-cell molecular heterogeneity at early embryonic stages, I selected CE as my separation technique because of its scalability in sample size. Due to the above-mentioned advantages, CE-MS has been applied for single-cell metabolomics[6, 26, 43, 49-50] and proteomics[26, 51-53] in the *X. laevis* model.

1.3 Multi-omics for a holistic picture of biological processes

Current research generally focuses on one type of biomolecule. This single 'omic strategy has enabled many breakthroughs in biomarker discovery[54-56] and cell and developmental biology[7, 43, 57-59]. However, single 'omic analysis faces challenges when studying biological processes in the context of complex biological systems.[35] This is because the single 'omic analysis lacks the capability to provide a big picture of multifaceted molecular networks[35], especially for regulatory processes which involve biomolecules at different levels. For example, the Krebs cycle (also called the citric acid cycle, TCA cycle, see **Fig. 1.4**[60]) is a process that oxidizes acetyl coenzyme A (acetyl-CoA) to generate energy for cells. In the Krebs cycle, both metabolites including citrate, malate, and succinate, and proteins including citrate synthase and succinyl-CoA synthetase are correlated to play indispensable roles to maintain the energy supply for cells. Moreover, the Krebs cycle is closely related to other energy production pathways such as fatty acids metabolism and glycolysis (see **Fig. 1.4**).[60] Therefore, to study diseases related to the Krebs cycle, it will be difficult to provide a holistic view from single omics only focusing on either metabolites or proteins.

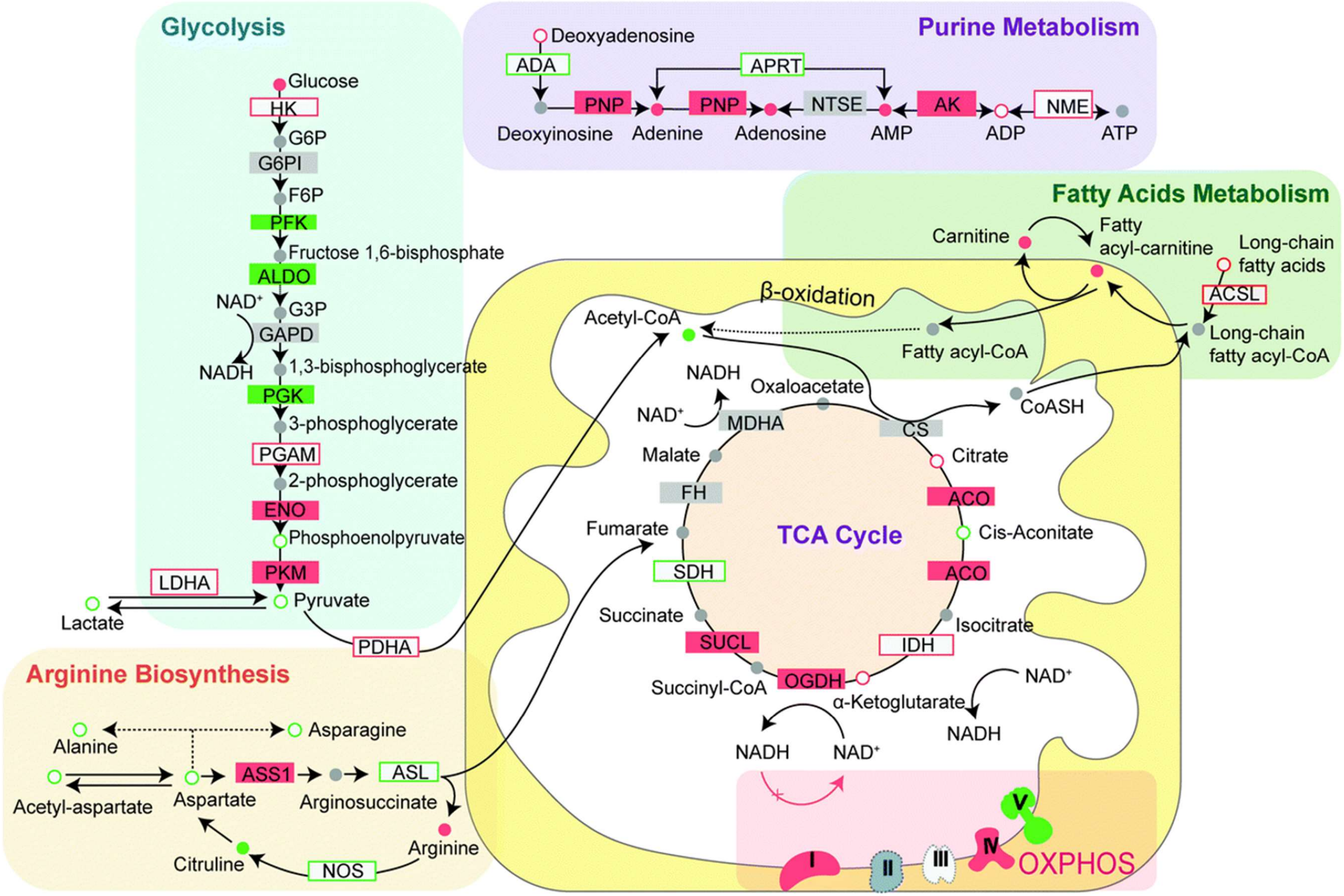


Figure 1.4 Krebs (TCA) cycle and its interaction with glycolysis, fatty acids metabolism, purine metabolism, and arginine biosynthesis. Reproduced from ref. 60 with permission from the Royal Society of Chemistry.

The integration of multiple ‘omics information deepens our understanding of the cells.[61-63] These different ‘omics can complement each other to filter out false positive or negative results. The multi-omics strategy has been applied for single-cell analysis[62, 64], disease research[65-66], and microbial biology[67-68]. For example, a previous study showed there was no difference in mRNA level across the dorsal-ventral axis for 8-cell *X. laevis* embryos according to single-cell transcriptomics of mRNA.[69] However, my dissertation research of the 8-cell *X. laevis* embryos using single-cell multi-omics (metabolomics and proteomics) found the left dorsal-animal (L-D1) clustered separately in the heatmap from the left ventral-animal (L-V1) cells,[26] indicating that there are cell-to-cell differences across the dorsal-ventral axis at such an early embryonic stage. Another example is that the multi-omics strategy advanced mitochondrial disease research.[70] By integrating transcriptomics, proteomics, and metabolomics of mammalian cells, transcription factor 4 (ATF4) was identified to play an active role in regulating the response to mitochondria stress in mammalian cells.[70] Furthermore, the multi-omics strategy was applied to study the growth of *E. coli* cells in response to different carbon sources.[13] The metabolomics dataset identified dysregulated metabolic pathways including the glycolysis I pathway with 50% of metabolites showing significant changes. Additionally, the integration of proteomics and transcriptomics datasets further confirmed the dysregulation of the glycolysis I pathway by revealing 67% gene changes and 39% protein changes. The overlapping genes, proteins, and metabolites are all significantly dysregulated, which improves the confidence of identified metabolic pathway changes involved in carbon source change.[13] These examples demonstrate the versatility of the multi-omics strategies for addressing complex biological questions and obtaining new knowledge.

1.4 Research model, significance, and motivations

My research in the Nemes Research Group used the South African clawed frog called *X. laevis* as the research model for cell and developmental biology studies. *X. laevis* provide biological and experimental advantages to my work as a vertebrate model. Firstly, *X. laevis* is close to humans in evolutionary history; therefore, the research findings based on *X. laevis* can be potentially applicable to humans.[71] Secondly, early embryonic cells of *X. laevis* embryos have known tissue fates.[72-73] For example, the midline dorsal-animal cell (D11, **Fig. 1.5**) develops into neural tissues, while the midline ventral-animal cell (V11, **Fig. 1.5**) develops into the epidermis.[72] The established cell fate map enables tracking the development of tissue-specific cell clones towards tissue during cell differentiation. Thirdly, *X. laevis* embryos develop externally to female frogs and have relatively large embryonic cell size (~ 250 µm in diameter at the 16-cell stage), which enables direct manipulations[74] such as microsampling and microinjection. Moreover, *X. laevis* embryos develop rapidly[75], which can shorten the experiment time. Additionally, the natural mating of *X. laevis* produces a large number of eggs (hundreds of eggs), which provides sufficient samples for research.

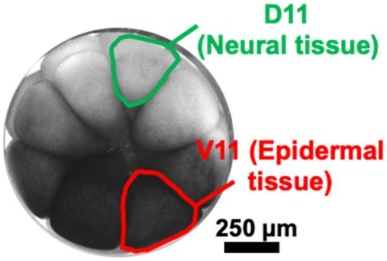
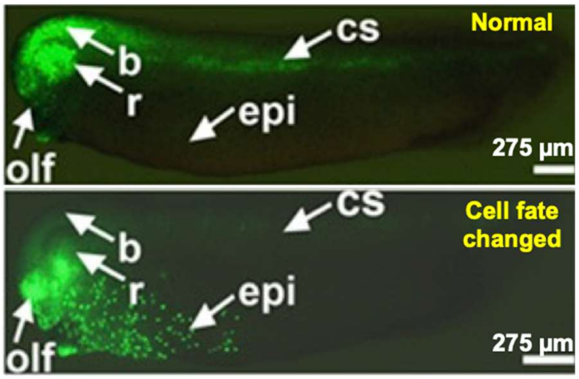


Figure 1.5 D11 and V11 cells in a 16-cell *X. laevis* embryo.

Due to the above-mentioned advantages, I used *X. laevis* to study the molecular composition and drivers of tissue-specific differentiation. To uncover the molecular mechanism underlying tissue formation, it was crucial for my work to be able to understand the molecular differences among single cells at the early embryonic stage with different tissue fates and explore how these molecular differences contribute to the respective tissues during embryonic development. Metabolites are active players in mediating cellular functions[76] and determining cell fate[77]. Therefore, investigating the roles of metabolites would help to uncover the mechanisms that guide single embryonic cells to form tissues during development. Our lab recently discovered a phenomenon that metabolites can alter a cell's tissue fate.[6] Acetylcholine and/or methionine, enriched in V11 cells, were demonstrated to induce cell fate change of D11 cells from neural to epidermal tissues (**Fig. 1.6**).[6] The discovered phenomenon underpinned the role of small molecules in cell fate determination. However, the molecular mechanisms underlying this cell fate change actions are unclear and are yet to be uncovered. Moreover, developing a technique that is amenable to isolating tissue-specific cells with varied cell sizes at different embryonic stages is necessary to explore the temporal molecular changes during cell differentiation and advance our understanding of tissue formation.



Note: brain (b), retina (r), epidermis (epi), olfactory (olf), and central somite (cs)

Figure 1.6 Metabolite-induced cell fate change. D11 cell fate changed from neural (top) to epidermal tissues (bottom). Reproduced from ref. 6.

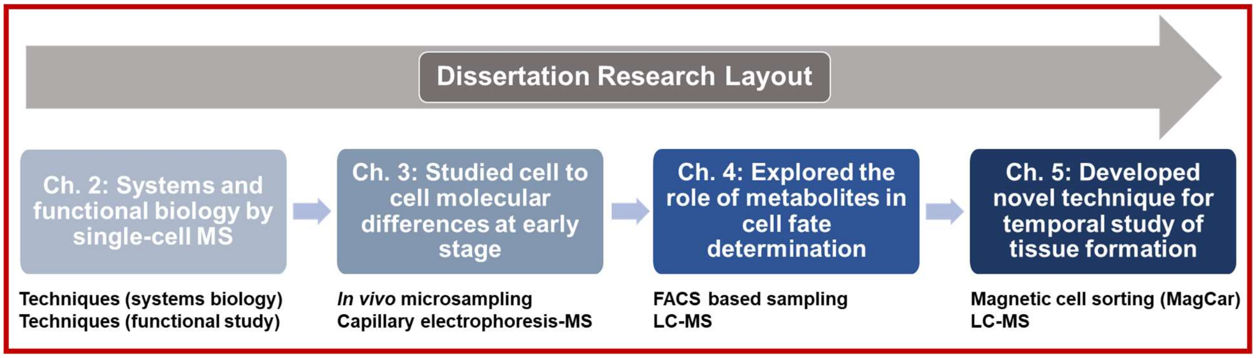


Figure 1.7 Overview of the dissertation research layout.

In this dissertation work, I developed techniques to answer a fundamentally important question: what are the molecular mechanisms that guide single cells to differentiate into their reproducible tissue fates during early embryogenesis? The dissertation research layout is shown in **Fig. 1.7**. MS combined with separation techniques to detect and identify metabolites and proteins was summarized in **Chapter 2**, which also contained protocols that we developed and validated for systems and functional biology.[78] Next, in **Chapter 3**, a workflow combining *in vivo* microsampling technique with single-cell MS was developed to enable a dual analysis of the proteome and metabolome of single embryonic cells while also providing anatomical and functional information in live, freely developing chordate embryos.[26] In **Chapter 4**, I explored the role of metabolites in driving cells' tissue fate during cell differentiation by developing fluorescence-activated cell sorting (FACS)-based sampling workflow for tissue-specific cell isolation. Along with LC-HRMS-based metabolomics and proteomics, I explored mechanisms underlying metabolite-induced cell fate change. In **Chapter 5**, to address the unmet sampling challenges for the temporal study of tissue formation, I developed a novel sampling technique called MagCar that allowed for the efficient isolation of tissue-specific cells from developing embryos and was also compatible with MS-based multi-omics. Finally, I summarized the results of my dissertation research and discussed the various areas that can be further explored to advance research on understanding tissue formation (**Chapter 6**). The obtained knowledge in my dissertation research could potentially be used to study healthy/dysfunctional tissue development, understand tissue regeneration, and improve our understanding of the causes of diseases, provided we know the molecular mechanisms of how these biomolecules guide single embryonic cells to develop into tissues.

Chapter 2: Systems and functional biology enabled by single-cell mass spectrometry analysis of metabolites and proteins

This chapter is reproduced and reused from a published book chapter: Erika P. Portero, Leena Pade, **Jie Li**, Sam B. Choi, and Peter Nemes*. Single-cell mass spectrometry of metabolites and proteins for systems and functional Biology, 2022, In Single Cell ‘Omics of Neuronal Cells. *Neuromethods* (NM, volume 184), pp. 87–114. Humana, New York, NY. Erika P. Portero and Leena Pade contributed equally with all other contributors. https://doi.org/10.1007/978-1-0716-2525-5_5. Reprinted/adapted by permission from Springer Nature. Copyright (2022).

My contributions involve in *in vivo* microsampling, sample processing for MS-based ‘omics (metabolomics workflow), LC-MS for metabolomic analysis and related data processing, and behavioral assays.

2.1 Abstract

Molecular composition is intricately intertwined with cellular function, and elucidation of this relationship is essential for understanding life processes and developing next-generational therapeutics. Technological innovations in capillary electrophoresis (CE) and liquid chromatography (LC) mass spectrometry (MS) provide previously unavailable insights into cellular biochemistry by allowing for the unbiased detection and quantification of molecules with high specificity. This chapter presents our validated protocols integrating ultrasensitive MS with classical tools of cell, developmental, and neurobiology to assess the biological function of important biomolecules. We use CE and LC-MS to measure hundreds of metabolites and thousands of proteins in single cells or limited populations of tissues in chordate embryos and

mammalian neurons, revealing molecular heterogeneity between identified cells. By pairing microinjection and optical microscopy, we demonstrate cell lineage tracing and test the roles the dysregulated molecules play in the formation and maintenance of cell heterogeneity and tissue specification in frog embryos (*X. laevis*). Electrophysiology extends our workflows to characterize neuronal activity in sections of mammalian brain tissues. The information obtained from these studies mutually strengthens chemistry and biology and highlights the importance of interdisciplinary research to advance basic knowledge and translational applications forward.

2.2 Introduction

Modern 'omics provide an unprecedented set of technologies to study connections between molecular composition and biological function. By enabling the detection and quantification of genes, transcripts, proteins, peptides, and metabolites, valuable information is gained on the molecular underpinnings of biological processes. High-throughput sequencing routine profiling of gene expression to large populations of single cells.[79-81] Identification and quantification of molecular markers from discovery 'omics help develop hypotheses and design experiments to evaluate their biological function.

Functional biological experiments leverage diverse types of approaches and technologies. Functional tests borrow tools of molecular biology to knock down or out genes, transcripts, or proteins, e.g., by using transcription or translation-blocking morpholinos and CRISPR-Cas9. Electrophysiology allows for eavesdropping on neurons. Different types of microscopies, typically optical to electron, enable the characterization of cell morphology, phenotype, and anatomy (**Fig. 2.1**). Single-cell RNA-sequencing with single-molecule fluorescence *in situ* hybridization (sm-FISH) recently captured molecular cell heterogeneity in the brain in high spatial and molecular resolution[82], supplementing classical knowledge of brain anatomy with molecular information[83]. By comparing single thalamic neurons projecting to motor, somatosensory, and visual cortices in the mouse brain, this approach uncovered several cell types within each projection. Technologies from modern 'omics empower classical biology and neuroscience with new investigative capabilities.

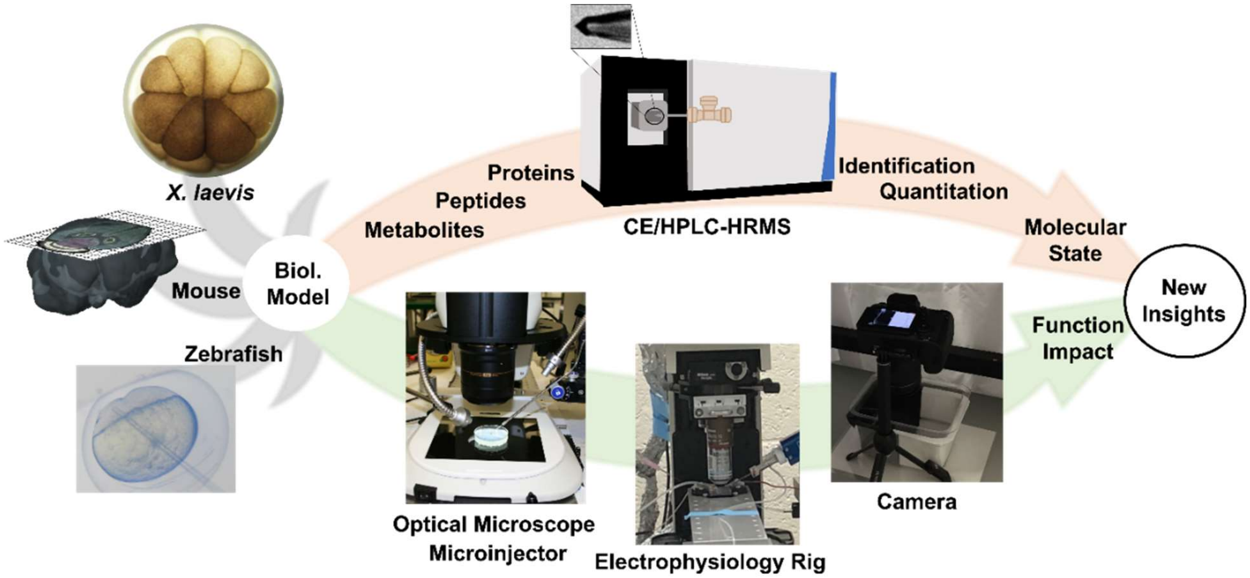


Figure 2.1 The approach to assess chemistry and function in biological models.The protocols discussed in this chapter were developed and validated using embryos from the frog *X. laevis* and zebrafish and sections from the mouse brain.

Single-cell mass spectrometry (MS) supports systems biological studies with information on the molecular state of cells. It complements single-cell transcriptomics by directly measuring proteomic, peptidomic, and metabolomic composition (**Fig. 2.1**). Detection with excellent molecular specificity facilitates identification without the requirement for functional probes. For example, MS does not necessitate antibodies for detection. High sensitivity and a broad linear dynamic range permit quantification of molecules at endogenous concentrations. Stringent reporting guidelines[84-85] and public repositories hosting data for reuse, reanalysis, and exchange (e.g., PRIDE[86] and MetabolomicsWorkbench[87]) promote scientific rigor and accountability in MS-based research. The current state of the field of single-cell MS was the focus of several reviews covering technology and application[29, 88-96]. Single-cell MS is adaptable to broad types of molecules, can be made sufficiently sensitive to quantify physiological concentrations, and is compatible with cells of broad dimensions and types, as well as different model systems used in biology and health studies.

Optical microscopy with single-cell MS integrates morphological and molecular information. Recent reviews provide a comprehensive discussion of single-cell MS[58, 90, 95-96]. **Figure 2.1** illustrates our protocol for optically guided single-cell MS in embryos of the South African clawed frog (*X. laevis*) and zebrafish and sections of the mouse brain. Cells were identified, dissected[6, 30] or their contents directly microaspirated[43, 49, 52] for metabolomic and proteomic analysis using capillary electrophoresis (CE) or liquid chromatography (LC) electrospray ionization (ESI) high-resolution MS (HRMS) (reviewed in[53, 97-99]). Microscopy with single-cell MS enabled detection of ~1,500 proteins in 2- to 50-cell *X. laevis* embryos and orthogonal validation using immunohistochemistry[100-101]. Single-cell proteomics by MS (SCoPE MS) quantified 3,000+ proteins from 1,490 cells[102-103], and a single-cell printer with liquid vortex capture enabled rapid metabolomics (~25 cells/min)[104]. With low attomole sensitivity, HPLC columns of narrow bore dimension or with a stationary phase supported on a monolith or a porous layer open tubular (PLOT) format allowed for the identification of ~1,300–4,000 proteins from 50–200 cells via magnetophoretic isolation from whole blood[105-107]. Nanodroplet processing in one pot for trace samples (NanoPOTS) identified over ~1,500 proteins from 10 HeLa cells and ~2,400 proteins from 100 pancreatic islet cells, supporting profiling across clinical samples[108]. These and other leaps in technology expanded the classical toolbox of cell biology, as was discussed in our recent review of the fields[53, 97-99].

We and others built CE-MS platforms to study biomolecules and their role in cell and neurobiological processes. CE combines advantages for single-cell analyses. The physical dimensions of fused silica capillaries are amenable to the limited amounts of samples that are contained in single cells. CE provides several methods for concentration enrichment in the capillary to boost sensitivity to low-abundance molecules (e.g., reviewed in[109]). An exquisite separation power and various data alignment strategies permit reproducible identifications[41, 110]. CE-ESI interfaces offer various designs to help hyphenate CE with HRMS for sensitivity, robustness, and reproducibility[111-112].

These custom-built CE-ESI-HRMS platforms revealed previously unknown details on cellular biochemistry. Proteins or metabolites were measured in single cells of *X. laevis* and zebrafish embryos (**Fig. 2.1**)[6, 30, 43, 52], single neurons dissected from *Aplysia californica*[27, 113] and mouse[41, 114], and single HeLa[115] cells. Our CE-ESI interface enabled the identification of hundreds of cationic and anionic metabolites[116] and ~700 proteins from ~5 ng protein digest from single *X. laevis* cells[52]. Targeted neuropeptides were detected in record sensitivity in the subfornical organ (SFO) and the paraventricular nucleus (PVN) of the mouse hypothalamus[117]. With a 200-zmol lower limit of detection, this technology also identified ~500 proteins from ~1 ng and ~225 proteins from ~500 pg protein digest, which estimates to a single neuron[8, 114, 118]. Our second-generation CE-ESI HRMS design employing a microprobe capillary enabled the *in situ* and *in vivo* analysis of single identified cells in live embryos[26, 43, 52] and the mouse brain[8, 119-120]. Its integration with cell labeling and stereomicroscopy permitted the tracing of tissue developmental trajectory. These instrumental capabilities revealed differences in the proteomic[22, 30, 52] and metabolic[6, 43] state of cells in *X. laevis* and zebrafish embryos, including those occupying the dorsal-ventral[6, 30, 121], animal-vegetal[51-52], and left-right[122-123] developmental axes in the frog. They also led to the discovery of metabolite-induced cell fate changes[6, 124] and metabolic communication between neighboring cells in *X. laevis*[125]. Further, the approach can be extended to patch-clamp electrophysiology, permitting the metabolomic[126] and proteomic[119-120, 127] characterization of identified neurons in the mouse brain. Most recently, we also integrated single-cell CE MS with assays measuring background color preference and swim pattern to assess the impact of metabolic perturbation on organismal behavior[26, 125].

This chapter presents our protocol enabling functional biological studies with insights into the proteomic and metabolomic state of cells in chordate embryos and mammalian neurons (**Fig. 2.1**). We overview required consumables and instruments (see ***Materials*** and ***Methods***) and discuss the experimental workflow. After identifying cells based on optical microscopy or electrophysiology and cell sampling by dissection or microprobe aspiration, the collected material is processed, and the resulting metabolites, peptides, or proteins are measured using LC and CE. Statistical analysis of signal abundances detected by HRMS–MS/MS allowed us to identify compounds for biological investigations. As an example, we present our approaches to test the effect of select compounds on tissue development. Representative examples are discussed with references to data showing the integration of single-cell 'omics with functional biology. The ***Note*** section advises on troubleshooting from the vantage point of an experimentalist, thus hoping to promote the combined use of single-cell HRMS-based proteomics/metabolomics with functional biology in other cell types and biological models.

2.3 Materials

2.3.1 Culturing embryos and neurons

Animals: Adult male and female *X. laevis* frogs (e.g., from Nasco, Fort Atkinson, WI); adult male and female zebrafish (e.g., from Zebrafish International Resource Center, Eugene, OR); Adult male mouse (e.g., from Charles River Laboratories, Wilmington, MA) (**see Note 1**).

Equipment: Incubators set to 14 °C and 18 °C; Stereomicroscope.

Solutions: Dejellying solution (2% cysteine): Dissolve 4 g of cysteine in 200 mL deionized (DI) water. Add 10 M sodium hydroxide dropwise to adjust pH to 8; 100% Steinberg's solution (SS): Mix 3.4 g sodium chloride (NaCl), 0.05 g potassium chloride (KCl), 0.08 g calcium nitrate ($Ca(NO_3)_2 \times 4 H_2O$), 0.205 g magnesium sulfate ($MgSO_4 \times 7 H_2O$), 0.66 g Trizma hydrochloride, 0.075 g Trizma base in 1 L of DI water. Adjust the pH to 7.4 by adding Trizma base and autoclave the solution. Store at 4–14 °C; Anesthetic solution (ketamine, 20 mg/mL and dexmedetomidine, 0.1 mg/mL): Prepare by mixing 200 µL of ketamine stock solution (100 mg/mL) and 200 µL of dexmedetomidine stock solution (0.5 mg/mL) with 800 µL of 0.9% saline solution for injection (e.g., part no. NDC0409-4888-06, Hopira, Inc., Lake Forest, IL); HEPES ringer solution: Prepare by mixing the following reagents to the following final concentrations: 86 mM NaCl, 2.5 mM KCl, 1.2 mM sodium phosphate (NaH_2PO_4), 35 mM sodium bicarbonate ($NaHCO_3$), 20 mM HEPES, 25 mM glucose, 5 mM sodium ascorbate, 2 mM thiourea, 3 mM sodium pyruvate, 1 mM $MgSO_4$, 2 mM calcium chloride ($CaCl_2$); Perfusion solution for mouse brain slices: Prepare by mixing the following reagents to final concentration: 126 mM NaCl, 21.4 mM $NaHCO_3$, 2.5 mM KCl, 1.2 mM NaH_2PO_4, 2.4 mM $CaCl_2$, 1.0 mM $MgSO_4$, and 11.1 mM glucose.

2.3.2 Embryology

Equipment:

Stereomicroscope for embryology to identify, inject, and microsample single cells (e.g., SMZ1270, Nikon, Melville, NY) and phenotype embryos, larvae, and tadpoles (e.g., SMZ18, Nikon). Microinjector to inject into or aspirate contents from identified cells (e.g., PLI-100A, Warner Instruments, Hamden, CT). Micromanipulator (e.g., MM-33, Warner Instruments). Capillary pullers (e.g., P-1000 for fused silica capillaries and P-2000 for borosilicate capillaries, Sutter Instruments, Novato, CA). Ancillary equipment: Heat block; Centrifuge (e.g., Refrigerated 5430R, Eppendorf).

Materials and Solutions:

Fine sharp forceps (e.g., Dumont #5, Fine Science Tools); Centrifuge vials (e.g., 0.2–0.5 mL LoBind tubes, Eppendorf); Borosilicate glass capillary (e.g., 0.5/1.0 mm inner/outer diameter, Sutter Instruments); Hair loop: Place both ends of a fine hair (~10 cm length) into a 6-inch Pasteur pipet to form a 2 mm loop. Secure it in place with melted paraffin. Sterilize before use by dipping in 70% ethanol and air-drying; Dissection dish: Cover the bottom of a Petri dish (35- or 60-mm) with non-toxic modeling clay (e.g., Claytoon was tested in *Xenopus* labs[128]). Make 3–5 wells of ~1.5 mm diameter using a cool glass bead to hold the embryos in place; Injection dish: Cover the bottom of a Petri dish (35- or 60-mm) with non-toxic modeling clay. Make several ~1.5 mm wells using a cool glass bead across the dish to hold ~20–30 embryos.

50% Steinberg's solution: Dilute 500 mL of 100% SS with 500 mL of DI water; 20% Steinberg's solution: Dilute 200 mL of 100% SS with 800 mL of DI water.

2.3.3 Sampling process

Equipment: Vacuum concentrator (e.g., CentriVap, LabConco, Kansas City, MO)

Reagents: All reagents are LC-MS grade to reduce chemical interference during MS detection. Methanol, anhydrous acetonitrile (ACN), water, acetic acid, 100 mM TEAB, 5% hydroxylamine, isobaric labeling kit (e.g., TMTsixplex, ThermoFisher Scientific or iTRAQ, AB Sciex).

Solutions: Metabolite extraction solvent: 50% (v/v) methanol in water containing 0.5% acetic acid; Proteomic digestion buffer: 50 mM ammonium bicarbonate containing protease inhibitor (1 protease inhibitor cocktail tablet per 10 mL); Trypsin (0.5 μg/μL) in 50 mM acetic acid; Patch-clamp solution: We use 50 mM ammonium bicarbonate in water as a compromise between sensitivity and function.

2.3.4 Mass spectrometry

Instrument and materials: CE system (e.g., laboratory-built following[6, 41] or CESI, AB Sciex, Toronto, Canada); High-resolution tandem mass spectrometer (e.g., quadrupole time-of-flight, Impact HD or timsTOF, Bruker Scientific, Billerica, MA or quadrupole orbitrap Q-Exactive Plus or Orbitrap Fusion Lumos, Thermo Scientific, **Fig. 2.1**); HPLC (e.g., Acquity I-class UPLC, Waters, Milford, MA and Dionex Ultimate 3000, ThermoScientific); Separation CE capillary (e.g., 40/100 µm inner/outer diameter fused silica, Polymicro Technologies, Phoenix, AZ); LC column (e.g., Acquity UPLC BEH Amide Column, 1.7 µm, 1 mm × 100 mm, and Acclaim PepMap C18 column, 3 µm, 0.075mm × 250 mm, Waters).

Solutions: All solvents and reagents are LC-MS grade; CE background electrolyte solution (BGE): 1% (v/v) formic acid in water for metabolomics; 25% (v/v) acetonitrile in water with 1 M formic acid for proteomics; CE-ESI sheath solution: 0.1% (v/v) formic acid and 50% methanol for metabolomics; 10% (v/v) acetonitrile in water with 0.05% acetic acid for proteomics; LC mobile phase for metabolomics using hydrophilic interaction LC (HILIC): For cationic separation, mobile phase A is aqueous 0.1% formic acid and B is acetonitrile containing 0.1% formic acid; for anionic separation, mobile phase A is aqueous 5% acetonitrile with 10 mM ammonium bicarbonate (pH 9) and B is aqueous 95% acetonitrile with 10 mM ammonium bicarbonate (pH 9); LC mobile phase for proteomics using reversed-phase LC (RPLC): Mobile phase A is aqueous 0.1% formic acid and B is acetonitrile containing 0.1% formic acid.

2.3.5 Functional studies

Instrument and Equipment: Epifluorescence stereomicroscope (e.g., SMZ18, Nikon, **Fig. 2.1**); Inverted microscope (e.g., Eclipse Ti-U, Nikon); Microinjector (e.g., PLI-100A, Warner Instruments); Micromanipulator (e.g., MM-33, Warner Instruments); Camera with tripod (e.g., ESO70D, Canon, **Fig. 2.1**); Software for processing movies (e.g., Windows media player software); Ancillary: Incubator set to 14 °C; Nutator rotator; Patch amplifier for electrophysiology (e.g., Sutter Instruments).

Materials: Tadpole food (e.g., part no. SA05964, Nasco); Transfer pipets; Black electrical tape; 26 G needle (e.g., part no. BD305115, Fisher Scientific); 1/2-gallon tank (e.g., part no. SB19271M, Nasco); Fine sharp forceps (e.g., Dumont #5); Inoculating turntable (e.g., part no. 50809-022, VWR).

Reagents: 200 proof ethanol; Gentamicin antibiotic (e.g., part no. 17-528Z, Fisher Scientific); Sylgard 184 silicone elastomer (e.g., part no. NC9285739, Fisher Scientific); Ficoll 400 (Sigma-Aldrich, St. Louis, MO); Benzocaine (part no. E1501-500G, Sigma-Aldrich); 1× phosphate-buffered saline (PBS) (Fisher Scientific); Fluorescent lineage tracer (e.g., fluorescent dextran 10,000 MW lysine fixable or mRNA lineage tracer, Invitrogen, Carlsbad, CA).

Solutions: 3% Ficoll in 100 % SS: Prepare by mixing 3 g of Ficoll in 100 mL of 100% SS; 4% paraformaldehyde: Prepare by mixing 4 g of paraformaldehyde in 40 mL DI water at 60 °C. Add a few drops of 1 N NaOH to adjust pH to 7.4. Add DI water to a total volume of 100 mL; DEPC water: Add 1 mL diethyl pyrocarbonate to 1 L DI water. Autoclave the solution, seal, and store at room temperature; 10% benzocaine: Weight out 2 g of benzocaine and place in a glass beaker. Add 20 mL 200 proof ethanol and stir using a magnetic bar in a stirrer; 2% benzocaine in 20% SS: Add 2 mL of 10% benzocaine to 8 mL of 20% SS dropwise, ensuring complete dissolution.

2.4 Method

This step lays out our methodology to culture *X. laevis* embryos to larvae or tadpoles and primary neurons from the mouse on the basis of established protocols[129-130]. *X. laevis* embryos require additional dejellying step for manipulation and functional experiments, as described below.

2.4.1 Frog embryos

- Obtain fertilized eggs by gonadotropin-induced natural mating of adult *X. laevis* frogs or *in vitro* fertilization as detailed elsewhere[129]. See **Note 1** on working with live vertebrate animals.
- Remove the jelly coat from fertilized eggs as follows:
    a) Remove excess media from dishes containing fertilized eggs.
    b) Add dejellying solution and keep embryos unperturbed for 2 min.
    c) Gently swirl dishes over a 2 min period and immediately decant excess dejellying solution.
    d) Transfer the embryos to a 250 mL beaker and add 10% SS. Gently swirl for ~30 s and decant excess liquid.
- Rinse embryos four times with 10% SS to remove remaining dejellying solution.
    a) Transfer ~300–500 embryos into individual 100 mm Petri dishes containing 100% SS. Place dishes in the 14 °C incubator (see **Note 2**).
- Collect two-cell embryos:
    a) Under a stereomicroscope, identify 2-cell embryos that display stereotypical pigmentation to accurately mark the dorsal-ventral axis, in reference to established cell fate maps[131].

b) Place the selected 2-cell embryos into a 100 mm Petri dish containing 100% SS and incubate at 14 °C until the desired developmental stage.

2.4.2 Brain section

Brain sections were collected following established protocols[132]:

- Anesthetize male mice aged postnatal day (PND) 21–35 with an intraperitoneal injection of anesthetic solution and perfuse with ice-cold HEPES ringer solution.
- After perfusion, dissect the brain rapidly in horizontal slices (220 μm) prepared in HEPES ringer solution using a vibratome.
- Recover slices for 1 h at 34°C in oxygenated HEPES holding solution. Then, place slide in the same solution at room temperature until use.

2.5 Sample collection from single cells

The goal of this step is to collect material from targeted single cells. In what follows, we give an example for *Xenopus.* The workflow starts with the identification of single cells in the *X. laevis* embryo in reference to established cell fate maps[72-73]. As shown in **Figure 2.2**, our laboratory established orthogonal strategies to collect single-cell samples. We dissected identified whole cells from *X. laevis* embryos or used fabricated microcapillaries as microprobes to aspirate portions of single identified cells from the embryo[43, 49, 53] or electrophysiologically characterized mouse neurons[119, 127].

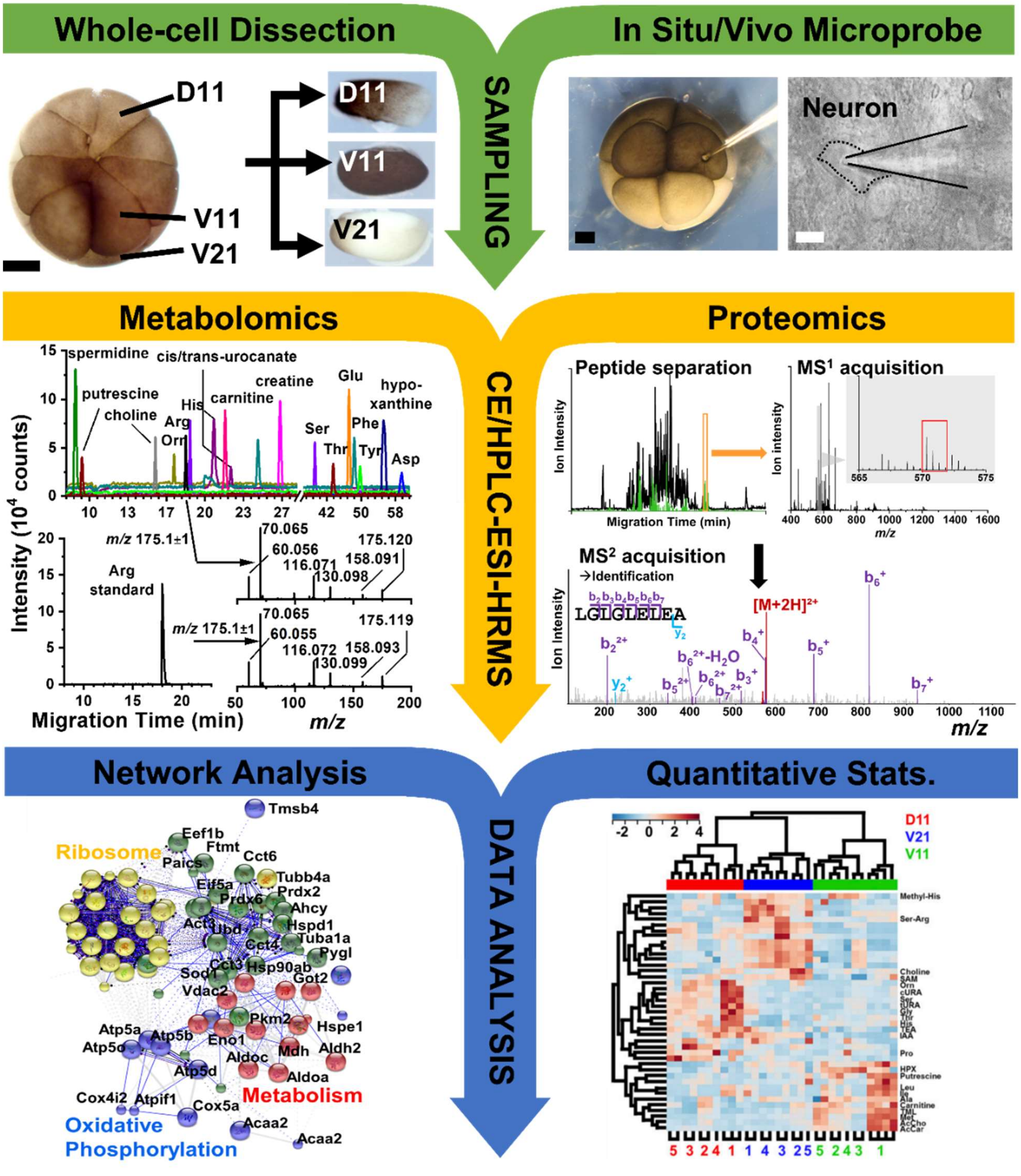


Figure 2.2 Our general single-cell MS workflow enabling the analysis of metabolites and proteins in single cells of *X. laevis* or zebrafish and single neurons in a section of the mouse brain. Scale bars: 250 µm (black), 20 µm (white).

2.5.1 Whole-cell dissection from X. laevis embryos

- Under a stereomicroscope, identify the cell of interest based on stereotypical cleavage and pigmentation, then dissect it manually as follows:
  a) Transfer the embryo into a dissection dish containing 50% SS.
  b) Using a hair loop, place the embryo of interest in a groove.
  c) Gently remove the vitelline membrane surrounding the embryo using sharp forceps.
  d) Use forceps to hold the embryo, preferably at the opposite side of the cell of interest. Lightly pull away the selected cell from the rest of the embryo.
- Transfer the isolated cell using a pipet into a LoBind Eppendorf vial containing chilled 10 µL methanol (~4 °C) or digestion buffer (see **Note 3**).
- Cool the Eppendorf vial (on ice) to preserve sample at low temperature and store samples at −80 °C until analysis and up to 3 months without detectable degradation.

2.5.2 In situ/vivo microsampling

- Fabricate the microprobe by pulling a borosilicate glass capillary to create a fine tip. We use a capillary puller (P-1000 Sutter Instrument) with custom settings: heat = 355, pull = 65, velocity = 80, time = 150.
- Using sharp forceps, cut the needle tip to an aperture of ~10–20 µm (see **Note 4**).
- Mount the microprobe into a capillary holder on a three-axis micromanipulator and connect its distal end to a microinjector.
- Transfer *X. laevis* embryos into an injection dish containing 50% SS. This protocol does not require removal of vitelline membrane.

- Use a stereomicroscope to aid viewing and manipulation of the embryo. Use a hair loop to immobilize the embryo of interest into a well in preparation for microsampling. Identify the cell of interest following protocols established elsewhere[72-73].
- Using a micromanipulator, guide the tip of the microprobe into the targeted cell to pierce through the membrane of the cell. Withdraw ~10–15 nL (or as needed) volume from the targeted cell by applying negative pressure to the microprobe using the connected microinjector.
- To end the microsampling, reduce pressure and retract the microprobe from the cell.
- Transfer the collected content into a LoBind Eppendorf vial containing 4 μL of metabolite extraction solvent or digestion buffer (see **Note 5**). We usually inject the collected material from the capillary by inserting the tip into the solvent/buffer and applying a positive pressure pulse.
- Cool the Eppendorf vial (on ice) to preserve sample at low temperature and immediately process the samples via metabolomics or proteomics workflows to prevent molecular degradation.

2.6 Sample Processing for MS-Based 'Omics

This section discusses protocols to process the collected materials for HRMS analysis. We use LC and CE to separate biomolecules in complementary performance prior to ESI-HRMS.

2.6.1 Metabolomics workflow

Extract metabolites from dissected single cells[6, 43] or aspirates collected by microprobe sampling[43, 116] as follows:

- For dissected single cells: Retrieve the single-cell samples stored in 100% methanol from –80 °C freezer and vacuum-dry them at 4 °C. Add 10 μL metabolite extraction solvent. Vortex-mix the vials for 30 s at room temperature to facilitate cell lysis and extraction of metabolites. Sonicate the sample vials for 3 min in an ice bath, followed by vortex-mixing for 1 min at room temperature.
- For microprobe aspirated samples: Retrieve single-cell samples collected in metabolite extraction solvent preserved on ice. Vortex-mix the vials for 1 min at room temperature to facilitate extraction of metabolites.
- For CE-ESI-HRMS, centrifuge the samples (dissected or microsamples) for 5 min at 8,000 ×g at 4 °C to pellet cellular debris. Proper pelleting is important to avoid CE capillaries from getting clogged. We usually store the aliquot together with the pelleted debris to avoid sample losses. The samples are kept at –80 °C until analysis.
- For HILIC-ESI-HRMS, centrifuge the samples at 13,000 ×g for 10 min at 4 °C and transfer the supernatant into a microvial and vacuum-dry the samples at 4 °C. Reconstitute the samples in 10 μL 95% (v/v) acetonitrile in water and centrifuge the samples at 13,000 ×*g* for 10 min at 4 °C to pellet potential debris that could clog the column (see **Note 6**). Transfer the supernatant into an LC vial and store the sample at –80 °C until analysis.

2.6.2 Proteomics workflow

- Lyse the collected cell or aspirate by sonication for 5 min. Heat the sample to 60 °C for ~15 min to denature proteins. We usually skip reduction and alkylation steps for our single cell samples for higher sensitivity[52].

- For one-step digestion, add ~50 ng of trypsin protease to the protein extract and incubate the mixture at 37 °C for 5–6 h. For neurons yielding less starting protein amounts, add ~2 ng trypsin and digest at 60 °C for 1 h.
- Vacuum-dry the resulting protein digest and store it at −80 °C until analysis.
- (Optional) To enable multiplexing relative quantification, barcode the dried protein digests. We use TMT isobaric labeling following the vendor's protocol (see **Note 7**), downscaled to the total amount of protein/peptide contained in the sample:
    a) Reconstitute the dried protein digest in 10 μL of 100 mM TEAB and tag it with 1 μL of 85 mM TMT reagent.
    b) Incubate each sample for 1 h at room temperature.
    c) Quench the reaction with 2 μL of 5% hydroxylamine and incubate the mixture for 15 min at room temperature.
    d) Mix the multiple tagged samples, vacuum-dry the mixture, and store it at −80 °C for up to 1 month until analysis.

High-resolution mass spectrometry (HRMS) for sample measurement

In this step, biomolecules in the resulting samples are separated and detected using ESI-HRMS. We present protocols for separation based on partition chromatography (LC) and electrophoresis (CE). These separation techniques provide complementary benefits in sensitivity, throughput, and molecular coverage (**Fig. 2.3**). The resulting data are processed using established approaches in bioinformatics, including but not limited to statistics, multivariate data analysis (e.g., principal component analysis and hierarchical cluster analysis), or machine learning (e.g., Trace[133-134]). These technologies and related protocols (**Fig. 2.2**) allowed us to document metabolic and proteomic differences between cells in embryos of *X. laevis* and zebrafish[6, 30, 52, 122] and single neurons in the mouse brain[119-120, 127].

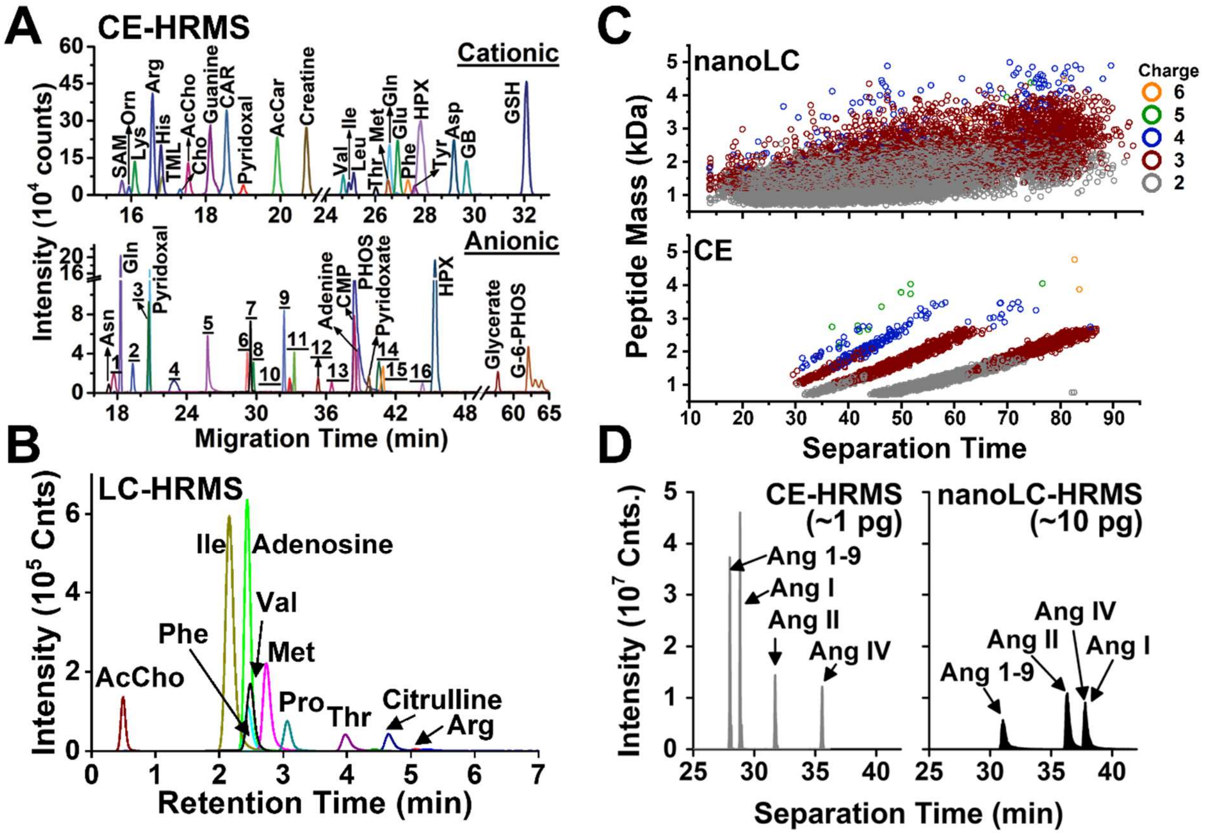


Figure 2.3 Representative detection of metabolites and proteins by CE and HPLC ESI-HRMS. (**A**) Chemical profiling of anionic and cationic metabolites in a single *X. laevis* cell using CE-ESI-MS. (**B**) HILIC LC-MS of polar metabolites from limited populations of cells. (**C**) Comparison of peptide identifications by CE-ESI-HRMS and nanoLC-nanoESI-HRMS. (**D**) Targeted detection of angiotensin peptides in the PVN and SFO of the mouse hypothalamus using CE and nanoLC.

2.6.3 CE-ESI-HRMS

- Construct the CE-ESI interface following protocols established by us and others[8, 41, 113, 135]. A simplified procedure to build a blunt-tip CE-ESI interface follows:

  a) Cleave a 1-m long piece of fused-silica CE capillary.

  b) On the outlet end of the CE capillary, burn off ~1.5 mm of polyimide coating and clean using isopropanol. Before proceeding to the next step, ensure that the capillary end is clean of burned residues to avoid the leaching of interfering ion signals from the burned residue.
  c) Feed the CE capillary outlet-end into a T-junction connected to a sheath-flow capillary.
  d) Mount the CE capillary into the T-junction to feed the CE capillary through the blunt tip emitter, allowing the capillary to protrude ~40–50 μm past the emitter.
- Hydrate the CE capillary by flushing with LC-MS grade water overnight.
- Position the tip of the CE-ESI interface ~5 mm from the inlet orifice of the mass spectrometer.
- Fill CE capillary with BGE and flush sheath-flow capillary with sheath solution.
- Initiate the electrospray as follows:
  a) Using a translation stage, fine-position the electrospray emitter tip ~2–3 mm from the mass spectrometer orifice to operate the electrospray in the cone-jet regime (ESI voltage 1.8–2.0 kV).
  b) Monitor the electrospray using a stereomicroscope.
  c) Observe the stability of total ion current (TIC) for ~10–15 min to ensure stable operation before analyzing a sample (see **Note 8**).
- Inject ~10–15 nL from the metabolite or protein extract hydrodynamically into the CE capillary following previously described protocols[41, 49].
- Gradually increase CE separation voltage from ground (0 V) to ~20–22 kV. Sudden application of high voltage may break the capillary.

- Load the MS method as described in **Table 2.1** (see CE-HRMS) and start data acquisition (**see Note 9**).

2.6.4 LC-ESI-HRMS

For Metabolomics:

- We use the following LC parameters: column temperature, 35 °C; autosampler temperature, 4 °C; injection volume, 1 µL; flow rate, 130 µL/min. Positive ion mode gradient: 0–0.5 min 95% B, 0.5–10 min 95–40% B, 10–13 min 40% B, 13–15 min 40–95% B, 15–22 min 95% B; Negative ion mode gradient: 0–0.5 min 99% B, 0.5–2.5 min 99–82.5% B, 2.5–6.5 min 82.5–68% B, 6.5–10 min 68–30% B, 10–13 min 30% B, 13–15 min 30–99% B, 15–22 min 99% B.
- Select MS method parameters described in **Table 2.1** (see LC-HRMS).

For Proteomics:

- LC parameters: 0–5 min 2% B, 5–85 min 2–35% B, 86–90 min 70% B, 91–120 min 2% B; autosampler temperature, 4 °C; injection volume, 1 µL; flow rate, 300 nL/min.
- Load the MS method parameters and start data acquisition. Our typical parameters are listed in **Table 2.1**. Adjust ion source settings to get a stable nanospray. (see **Note 8 and 9**).

2.7 Data processing

2.7.1 Metabolomics

- Survey the MS–MS/MS data for molecular features (signals with unique *m/z* and separation time) using available software packages. For example, we employ

MetaboScape Version 4.0.4 (Bruker Daltonics) using the following settings: intensity threshold, 1,000 counts; minimum peak length, 5 spectra.

- Annotate metabolites based on the accurate mass, isotopic distribution pattern, and tandem MS spectra against reference spectra available in MS-MS/MS databases, including but not limited to METLIN[136], EMBL (http://curatr.mcf.embl.de/), MzCloud (https://www.mzcloud.org/), MassBank of North America (https://mona.fiehnlab.ucdavis.edu/), and the Human Metabolome Database[137]. For example, we use METLIN with an annotation tolerance ≤ 10 ppm mass accuracy and MS/MS score ≥ 700–900.
- Perform relative/absolute quantification using under-the-peak-areas (label-free quantification) or ion signal abundances (multiplexing quantification) serving as a proxy for metabolite abundance.
- Perform statistics and multivariate data analysis to select molecules for follow-up functional studies (see **Fig. 2.2**).

2.7.2 Proteomics

- Identify proteins using established bioinformatics software packages broadly available for bottom-up proteomics. For example, we analyze the MS–MS/MS data in ProteomeDiscoverer (Thermo Fisher Scientific) or MaxQuant (Max Planck Institute of Biochemistry) against the mouse or *Xenopus* proteome (e.g., downloaded from UniProt[138] or Xenbase[75, 139]) with the following search parameters: digestion enzyme, trypsin; missed cleavages, up to 2; variable modification, methionine oxidation; precursor mass tolerance, 20 ppm; fragment mass tolerance, 4.5 ppm; minimum peptide length, 5. Peptides are filtered to <1% false discovery rate (FDR), calculated against a reversed-sequence decoy database. The reported proteins are grouped based on the closest parsimony principle. We remove common contaminants from the final list of protein identifications by manually annotating for common contaminant proteins (downloaded from UniProt).

- Employ label-free or label-based strategies and software packages from the proteomics community to compare protein levels between single cells. For example, we used MaxQuant Version 1.5.5.1[140] or Proteome Discoverer (Thermo Scientific) to quantify the proteomic state of single embryonic cells and neurons by calculating label-free quantitative indexes (LFQ values)[51-52] or relative reporter ion signal abundances from TMTs[30].

2.8 Functional studies

The goal of this step is to link chemistry with biological function. As an example, we describe a protocol to prepare brain tissues to record neuronal activity followed by *in vitro* single-cell proteomics[119, 127]. In the context of cell differentiation, we trace cell fates to understand how cells divide to form specific tissues and organs. Because cell fates are reproducible in *X. laevis*[141-142], it is possible to inject molecules into identified cells to determine their developmental impact on tissue specification and organogenesis[6, 124-125]. **Figure 2.4** shows an example, in which the vehicle or select metabolites were injected into specific cells, while fluorescently monitoring their tissue clone via the co-injection of the green fluorescent dextran. Alternatively, fluorescent proteins can be expressed in the cell, for example, by injecting the corresponding mRNA[6, 128]. In *X. laevis* tadpoles, it is also possible to perform behavioral assays to assess sensory (e.g., visual), muscular, cognitive, and other functions. As an example, we adopted the background color preference assay[143] to test the behavior of pre-metamorphic *X. laevis* tadpoles after performing *in vivo* single-cell MS on the precursor embryo[26].

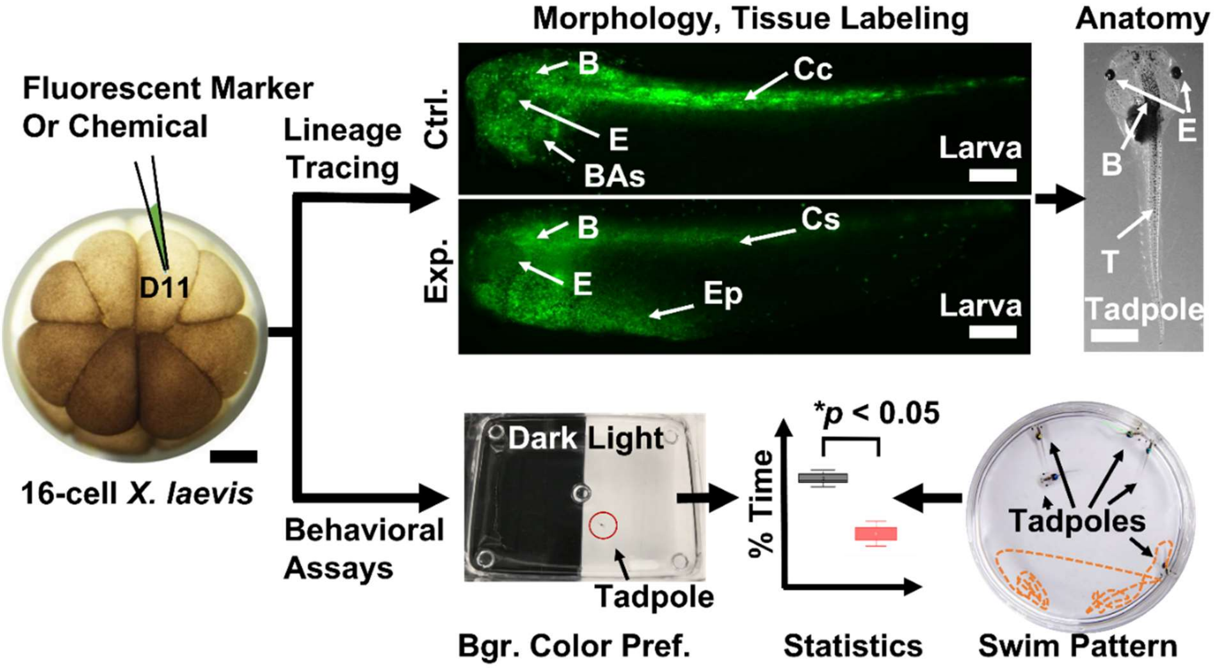


Figure 2.4 Techniques to investigate chemistry and function during development. (Top panel) Analysis of cell fate, morphology, and anatomy following fluorescence lineage tracing of the left dorsal-animal midline cell (D11) in control (Ctrl.) and experimental (Exp.) *X. laevis* larvae. (Bottom panel) Background color preference and swim pattern assays evaluating behavior in *X. laevis* tadpoles. Key: B, brain; BAs, branchial arches; Cc; central somites; E, eye; Ep; epidermis. Scale bars = 250 µm (embryo, larvae), 1.5 mm (tadpole).

2.8.1 Single-neuron electrophysiology and capillary microsampling

- Perfuse midbrain slices continuously at 1.5–2 mL/min with perfusion solution at 28–32 ºC, following established protocols[144].
- Backfill the patch pipettes for recording (2–4 MΩ) with ~20 µL of 50 mM ammonium bicarbonate in water (see **Note 10**).
- Putatively identify neuron type of interest. For example, dopaminergic neurons can be identified based on their location in the lateral portion of the substantia nigra and their size. Consult with brain anatomy atlases to improve the accuracy of tissue identifications.

- Neuron identification may be aided by electrophysiology. Detection of a slow pacemaker firing pattern (> 2 ms action potential), indicates dopaminergic neurons[145].
- Obtain a giga-ohm seal and record action potential for 60 s in a cell-attached configuration. A patch-clamped neuron is demonstrated in **Fig. 2.2**. We used the Sutterpatch software to control the devices[119-120, 127].
- Following electrophysical analysis, apply a steady negative pressure at the outlet end of pipette with a syringe to aspirate a portion of neuronal soma (see **Note 11**).
- Under an inverted microscope (40× magnification), visually inspect the neuron during microaspiration. A slight reduction in neural soma size is anticipated from successful patching. **Figure 2.2** shows a neuron during sampling after electrophysiological recording.
- Gently withdraw the pipette from the cell then expel the collected contents into an LoBind microvial containing 5 μL protein digestion solution chilled on ice. After a 1 h digestion at 60 ºC, store samples at −80 °C for up to 1 month without detectable degradation, until analysis.

2.8.2 *Cell lineage tracing*

- Prepare fluorescent tracer solution:
  a) Prepare a 100 μL solution containing 0.5% fluorescent dextran in DEPC water.
  b) (Optional) Synthesize capped mRNA using an *in vitro* transcription kit. Prepare a working solution containing ~50–100 pg/nL mRNA pellet in DEPC-treated water.
- Prepare injection needles following steps described earlier. Calibrate the volume of injection by injecting water droplet into mineral oil and measuring the diameter of the water droplet.

- Fill the injection needle with ~0.5–1 μL fluorescent tracer solution using a microinjector in "fill" mode following established protocols[128].
- Transfer the embryo into an injection dish containing 3% Ficoll in 100% SS.
- Under a stereomicroscope, use a hair loop to gently angle the embryo to orient the cell of interest for facile access for injection (**see Fig. 2.4**).
- Using a calibrated micromanipulator, guide the tip of the injection capillary into the targeted cell. Inject ~1–5 nL of the sample by applying +40 psi on the capillary for ~300–500 ms using a microinjector (see **Note 12**).
- Culture the injected embryos in 3% Ficoll in 100% SS for 3–4 h to allow the cell membrane to heal. Transfer the embryos to 50% SS and culture at 14–22 °C until the larval stages 32–34 (see **Note 13**).
- Anesthetize the larvae or tadpoles on ice or using 0.5% benzocaine in 100% SS. Ensure the success of anesthesia by gently touching the tadpole with a capillary, anticipating no response if successful. Increase the concentration of benzocaine if necessary. Ensure tadpoles are handled and treated humanely so that the organisms do not suffer or feel pain (see **Note 1**).
- Fix the specimens in 4% paraformaldehyde for 1 h on a rotator.
- Rinse the tadpoles with 1× PBS twice. Store the tadpoles in 1× PBS at 4 ºC.
- For imaging, mount the specimens in a 30 mm dish containing ~1–2 mL of 1× PBS. Image the specimens using epifluorescence microscopy (e.g., SMZ18, Nikon) (see **Note 14**).
- Acquire images using a microscope following the manufacturer's instructions.

- Conduct lineage analysis by determining the relative contribution of fluorescent cells to tissues and organs (see protocols in References[72-73, 141]).

2.8.3 Behavioral assay

Tadpole preparation

- Obtain and culture embryos as described earlier.
- Inject identified cells with the test compounds as described earlier. Prepare negative control by injecting identified cells with DEPC-treated water. Use non-injected embryos for later use as the wild-type control group.
- Place the embryos in the 14 °C incubator until they reach gastrula stage. Transfer the embryos to room temperature in 90 mm Petri dishes containing 20% SS and change media every two days and culture the tadpoles until the feeding stage (Stage 45) (see **Note 1**).
- Feed tadpoles every other day as follows:
  a) Mix food with 20% SS solution to form a paste.
  b) Place the paste in a corner of the Petri dish containing the tadpoles.
  c) Provide more food as needed.
- Maintain the tadpoles under a 12 h light/dark light cycle.

Background color preference assay

- The color preference assay is performed in a setup consisting of nested tanks following an established protocol[143]. The inner tank holds the tadpoles. The outer tank provides the background colors with half covered with a black tape and the other half covered with a white paper. Ensure both tanks are water-leveled to aid visual inspection of tadpole behavior.

- Fill the inner test tank with 20% SS to the 5 cm water mark from the top and insert the inner tank inside the outer tank.
- Mount the camera on a tripod to record the entire tank from above (see **Note 15**).
- Transfer a single tadpole on the white background of the inner tank. Limit this experiment to one tadpole at a time to avoid interactions between tadpoles which may confound behavioral phenotypes.
- Record the swim pattern of tadpole for 2 min.
- After 2 min, carefully lift the outer tank, rotate the outer tank, and return the inner tank into the outer tank. This step helps minimize the impact environmental factors may have on behavior. Start recording immediately and set the timer for 2 min.
- Place the tadpole back to a holding tank containing 20% SS. Record 2 trials for each of ~10–15 tadpoles.
- Repeat the assay on the next day on the same tadpoles to test for reproducibility and enhance statistical evaluation of the results by obtaining more data.

Euthanasia

- Place tadpoles in a 90 mm Petri dish containing 2% benzocaine in 20% SS for ~15–20 min.
- Monitor the tadpole's reflex by gently touching with a hair loop. Anesthetized tadpole cannot swim or respond to mechanical stimuli (gentle touching). Only proceed to the next step if the tadpole is anesthetized.
- Freeze the larvae/tadpoles at –20 ºC overnight.
- Dispose of the tadpoles following protocols approved by the relevant institutional and federal authorities (**see Note 1**)

2.9 Notes

Note 1. *X. laevis,* mouse, and zebrafish are sentient and vertebrate animals; therefore, protocols pertaining to the care and handling of the animals must be approved by institutional and federal agencies. The work presented in this chapter was approved by the Institutional Animal Care and Use Committee of the University of Maryland (approval numbers R-DEC-17-57, R-FEB-21-07, and R-JUN-20-31) and/or The George Washington University (approval numbers #A311 and #A283).

Note 2. Development is temperature-dependent in *Xenopus*[139, 146], thus providing a helpful tool to time biological and chemical experiments. Low temperatures slow down the speed of cell-cleavage, extending the time to select embryos for experiments.

Note 3. For metabolomics, place cells in ice-cold 100% methanol immediately upon collection. Methanol denatures enzymes and low temperatures slow down chemical reactions, thus minimizing metabolic changes.

Note 4. Capillaries with too large or too small diameters challenge microsampling or microinjection. In *Xenopus*, we find apertures larger than ~20 µm tend to cause substantial damage to the cell membrane. Without the membrane being able to heal, the cytoplasm may leak into the media and the cell may not be able to continue division for functional experiments. Conversely, apertures below ~10 µm may clog with yolk and cytoplasmic content, requiring refabrication of the microprobe.

Note 5. Tailor the composition of the metabolite extraction solvent to the type of metabolites of interest in a particular study. To study polar metabolites, we use aqueous 40% (v/v) acetonitrile with 40% (v/v) methanol as the extraction solution. Theoretical predictions based on partition and distribution coefficients can help experimental design[122].

Note 6. Debris may clog CE capillaries or LC columns. We find that centrifugation of samples before analysis prevents clogging. The efficiency of centrifugation depends on the field force and time of centrifugation. We typically use 13,000 ×g for 10 min.

Note 7. To select the spectral resolution appropriate for the multiplexing relative quantification, refer to instructions from the manufacturer of the reagents.

Note 8. We consider the electrospray to be stable when the total ion current exhibits less than ~15% relative standard deviation over ~40 min of separation.

Note 9. Tailor MS–MS/MS experimental settings to the chosen separation technology. For example, we adjust the number of targeted molecular features and the duration of dynamic exclusion depending on typical peak widths and the complexity of the sample in LC and CE experiments (see **Table 2.1**).

Note 10. It is imperative to optimize the composition of the intracellular solution used for patch clamp electrophysiology to the osmolarity of the neuron. Although potassium gluconate is commonly used in electrophysiological recordings, we use ammonium bicarbonate to minimize spectral and ionization interferences during MS caused by involatile salts.

Note 11. Aid sample collection by applying consistent negative pressure and continuously monitoring the size of the neuron under the inverted microscope. Shrinking neuronal soma is an indication of successful sampling. Be careful not to aspirate the media surrounding the neuron to avoid sample dilution and interferences due to salt during MS analysis.

Note 12. Limit damage to the cell membrane by carefully withdrawing the needle tip from the cell. With negligible damage, the cell membrane heals, and the embryo continues development. Damage to the cell membrane can cause leakage of cytoplasmic content, which can result in low survival rates or lethality. Take extra care to also avoid damaging the neighboring cells to facilitate development.

Note 13. Based on our experience and other protocols[142], culturing at lower temperature (14–16 °C) improves survival rates.

Note 14. Clearing agents (e.g., benzyl alcohol:benzyl benzoate, BA:BB) may be used in *X. laevis* embryos and tadpoles to improve fluorescent imaging[147]. Tissue clearing is recommended for imaging deep in intact embryos/tadpoles and sections.

Note 15. The locations of the eyes in tadpoles help accurately determine crossings between the white and black backgrounds (see **Fig. 2.4**). Therefore, the camera used to record video trials should have sufficient optical resolution and frame rate to clearly identify the eyes and monitor fast tadpole swimming. We typically use ~30 fps to monitor tadpoles during both the background color preference and swim assays.

Table 2.1 Our typical instrumental settings for detecting metabolites and proteins in CE and LC ESI-HRMS using the positive (+) and negative (–) ion mode.

Parameters	**CE-ESI-HRMS**		**LC-ESI-HRMS**	
Instrument Configuration	Q-TOF	Q-OT	Q-TOF	Q-OT
Compounds	Metabolites	Proteins	Metabolites	Proteins
MS survey scan frequency	MS^1 –MS^2, 2 Hz	MS^1, 7; MS^2, 13 Hz	MS^1 –MS^2, 2 Hz	MS^1, 7; MS^2, 13 Hz
Mass range (*m/z*) and spectral resolution	50–550 at 40,000 FWHM	400–1,700 at 35,000 FWHM for MS^1 and 17,500 FWHM for MS^2	50–1,300 at 40,000 FWHM	400–1,700 at 35,000 FWHM for MS^1 and 17,500 FWHM for MS^2
Dry gas	2 L/min (N_2)	–	4 L/min (N_2)	–
Nebulizer gas	–	–	0.4 bar (+); 1 bar (–)	–
Dry temperature	100 °C (+); 150 °C (–)	275 °C	220 °C	275 °C
Ion-transfer capillary voltage	–1,700 V (+); +2,100 V (–)	2,300 V	–4,500 V (+); +4,000 V (–)	1,800 V
Collision energy	18 eV (CID)	36% (normalized, HCD)	15–35 eV (CID)	36% (normalized, HCD)
Isolation *m/z* window	1.5 Da	1.5 Da	1.5 Da	1.5 Da
Dynamic exclusion mass tolerance	5.0 ppm	5.0 ppm	5.0 ppm	5.0 ppm
Dynamic exclusion	9 s	9 s	13 s	13 s
High intensity ion signal threshold (counts)	1.5×10^6	1.5×10^6	1.5×10^6	1.5×10^6
AGC target (counts)	–	1×10^6	–	1×10^6
Minimum AGC target (ion counts)	–	9.2×10^2	–	9.2×10^2
Maximum injection time (ms)	–	50	–	50
DDA Top N	5	20	5	20

Chapter 3: Single-cell mass spectrometry with *in vivo* microsampling enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*)

This chapter is reproduced and reused from a published peer-reviewed paper from: Camille Lombard-Banek[¶], **Jie Li**[¶], Erika P. Portero, Rosemary M. Onjiko, Chase D. Singer, David Plotnick, Reem Q. Al Shabeeb, and Peter Nemes*, *In vivo* subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*), *Angewandte Chemie International Edition*, 2021, 60 (23), 12852–12858. [¶]Co-first author. https://doi.org/10.1002/anie.202100923 with permission from the publisher John Wiley and Sons.

My contributions (the lead co-first author): I designed the research; I prepared/measured the metabolomic samples; I designed/conducted the survival, anatomical, and behavioral assays; I analyzed the data, interpreted the results, and wrote the manuscript.

3.1 Abstract

We report the development of *in vivo* subcellular high-resolution mass spectrometry (HRMS) for proteo-metabolomic molecular systems biology in complex tissues. With light microscopy, we identified the left-dorsal and left-ventral animal cells in cleavage-stage non-sentient *X. laevis* embryos. Using precision-translated fabricated microcapillaries, the subcellular content of each cell was double-probed, each time swiftly (<5 s/event) aspirating <5% of cell volume (~10 nL). The proteins and metabolites were analyzed by home-built ultrasensitive capillary electrophoresis electrospray ionization employing Orbitrap or time-of-flight HRMS.

Label-free detection of ~150 metabolites (57 identified) and 738 proteins found proteo-metabolomic networks with differential quantitative activities between the cell types. With spatially and temporally scalable sampling, the technology preserved the integrity of the analyzed cells, the neighboring cells, and the embryo. 95% of the analyzed embryos developed into sentient tadpoles that were indistinguishable from their wild-type siblings based on anatomy and visual function in a background color preference assay.

3.2 Introduction

Unbiased measurement of transcripts, proteins, and metabolites in a live cell promises to revolutionize investigations in molecular systems cell biology. Even today, after the invention of single-cell transcriptomics[29, 90-91, 148], there exists no single technology capable of the unbiased characterization of both proteins and metabolites in the same single cell *in vivo* to enable studies at the level of the functioning organisms. While tools of molecular biology and high/super-resolution optical microscopy empowered systems biology for live organisms, they only work for a limited number of gene products at a time that also have to be known ahead of time, thus limiting research scope to targeted studies, typically building on prior knowledge.

Single-cell high-resolution mass spectrometry (HRMS) emerged as a powerful alternative for targeted and untargeted (discovery) studies (reviewed in refs.[29, 48, 90, 148-150]). Specialized technologies in cell handling, microscale sample processing[30, 49, 52, 102, 151-152], high-efficiency chemical separations[41, 105, 153], and ion generation extended time-of-flight and orbitrap HRMS instruments to sufficient sensitivity and throughput for single-cell analyses. For example, using nanoHPLC, NanoPOTS enabled the detection of 650+ proteins in single HeLa cells,[152] which was recently increased to 850+ proteins on a new-generation mass spectrometer.[105] ScoPE-MS also used nanoHPLC to quantify ~750 proteins from a mammalian cell.[102] The technology was recently advanced to the quantification of ~1,000 proteins/cell.[102, 154] MALDI-TOF followed by nanoLC-HRMS enabled the characterization of lipids, peptides, and proteins in large numbers of single cells dissociated from a rat root ganglion.[155] The data emerging from these and other single-cell studies can be used to derive new hypotheses for biological investigations to establish the functional implications of the observed molecular cell heterogeneity.

We and others custom-built ultrasensitive capillary electrophoresis (CE) electrospray ionization (ESI) platforms to enable single-cell analyses. These instruments were used to detect transcripts, ~3,000 proteins, or ~100 metabolites in single cells/neurons dissected from *Aplysia californica*[27, 41-42, 156], *X. laevis* embryos[6, 30, 51], or the rat[41]. Using capillary microsampling, we extended these measurements to the direct analysis of metabolites[43, 121] or proteins[52] in single cells in early developing *X. laevis* and zebrafish embryos. Although single-cell HRMS technologies usher an emerging era of multi-omic single-cell 'omics,[64, 157] these tools so far required isolation or sorting of the cells, preventing studies *in vivo*. We here report the first example of *in vivo* single-cell HRMS that enables dual proteo-metabolomics of spatiotemporally identified single cells in a live embryo, which freely develops to a normally behaving tadpole post analysis.

3.3 Experimental section

3.3.1 Materials and regents

All reagents and solvents were obtained at reagent grade or higher purity from Fisher Scientific (Pittsburg, PA) unless otherwise noted. Solutions for HRMS were prepared using LC-MS grade solvents and reagents (methanol, acetonitrile, water, formic acid, and acetic acid). For CE, bare fused silica capillaries (40/105 µm inner/outer diameter) were purchased from Polymicro Technologies (Phoenix, AZ) and used after initial conditioning (100 mM sodium hydroxide for 5 min, then rinsed with water). CE micro-flow ESI (CE-μESI) for metabolomics employed a laser-cleaved stainless-steel blunt needle as the spray emitter (130/260 μm inner/outer diameter) from Hamilton Company (Reno, NV). The CE-nanoESI setup for proteomics used a borosilicate capillary (0.75/1 mm inner/outer diameter) tapered on a Flaming/Brown-style capillary puller (P-1000, Sutter Instruments, Novato, CA) and cleaved to a 20-μm tip diameter under a stereomicroscope (model SMZ18, Nikon, Melville, NY).

3.3.2 Solutions

The *"metabolite extraction solvent"* was an aqueous mixture of 40% (v/v) acetonitrile (ACN) and 40% (v/v) methanol (MeOH) solution, chilled to 4 °C. The *"protein extraction–digestion buffer"* was 50 mM ammonium bicarbonate in water. For CE, the *"background electrolyte (BGE)"* was 1% (v/v) formic acid (yields pH 2.8[50]) for metabolomics and 25% (v/v) MeOH with 1 M formic acid for proteomics (yields pH 1.8). The *"CE-μESI sheath solution"* was aqueous 50% (v/v) methanol containing 0.1% (v/v) formic acid. The *"nanoESI sheath solution"* was 10% (v/v) MeOH with 0.05% (v/v) acetic acid.

3.3.3 Animal care and embryology

All protocols concerning the humane care and handling of vertebrate *X. laevis* animals were approved by the University of Maryland Institutional Animal Care and Use Committee (IACUC No. R-DEC-17-57). Adult male and female frogs were received from Nasco (Fort Atkinson, WI) and maintained in a breeding colony. Embryos were obtained by gonadotropin-induced natural mating of two sets of parents at ~21 °C, dejellied in 2% cysteine solution, and cultured in 100 % Steinberg's solution (14 °C) following standard protocols.[129] Two-cell embryos displaying stereotypical pigmentation, size, and location[129] were transferred into a Petri dish containing 100% Steinberg's solution. The embryos were monitored under a stereomicroscope until they reached the 8-cell stage, then the left dorsal-animal (L-D1) and left ventral-animal (L-V1) cells were identified based on reproducible cell-fate maps[158].

3.3.4 Microprobe sampling and sample processing

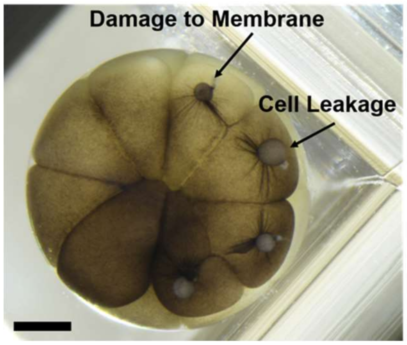


Figure 3.1 Microsampling was scaled to minimize cell damage. A 16-cell embryo with substantial damage after microsampling before technology refinement. Scale = 250 μm.

Contents of the L-D1 and L-V1 cells were collected via capillary microsampling[43, 49, 52, 121] in a randomized order. Microsampling using capillaries with too large tips or delaying aspiration may significantly damage the cell and the developing embryo. For example, in the 16-cell embryo shown in **Figure 3.1**, the cytoplasm of each cell partially leaked through the hole that was left on the cell membrane as the capillary with too large tip diameter was withdrawn. To promote cell and embryonic viability post microsampling, the diameter of the capillary tip and duration of sampling were scaled. In this study, the tip of a borosilicate capillary (0.75/1 mm inner/outer diameter) with ~20 μm diameter was fine positioned into the identified cells under control by a three-axis manual micromanipulator (Warner Instruments, Hamden, CT) and guidance by a high-resolution stereomicroscope (SMZ18, Nikon). An ~10 nL portion of the cell was aspirated each time by delivering calibrated pressure pulses (−40 psi) to the microcapillary using a connected microinjector (model PLI-100A, Warner Instruments, Hamden, CT). The total duration of each microsampling was less than ~3 s, separated by an ~30 s gap required to install the replacement capillary (for the second sampling). The subcellular aspirates were expelled into individual microvials and processed differently for metabolomic[49, 122] and proteomic[52] analyses following our established protocols. Metabolites were extracted into 4 μL of *metabolite extraction solvent* (chilled to 4 °C), facilitated by vortex-mixing for 1 min, and the resulting extract was centrifuged to pellet cell debris (8,000 ×*g*, 5 min, 4 °C). Proteins were extracted into 5 μL of *protein extraction–digestion buffer*, denatured at 60 °C, and digested to peptides with 0.5 μL of 0.5 μg/μL trypsin (37 °C, 5 h). The resulting digests were dried in vacuum at 60 °C. The resulting metabolite and protein samples were stored at −80 °C until analysis.

3.3.5 Behavioral assays

A portion of the microsampled (wild-type) 8-cell embryos was cultured to the tadpole stage (~21 °C) to study their visual behavior using a background c olor preference assay. This assay reproducibly and robustly measures the differential preference of tadpoles to spend time over light and dark backgrounds between Nieuwkoop-Faber Stages 45 and 50.[143] The tadpoles were humanely raised, maintained, and treated following protocols approved by the University of Maryland IACUC (protocol no. R-JUN-20-31). The residence times of each tadpole over the light (white) and dark (black) backgrounds were quantified in a double-tank behavioral system that we built and operated following an established protocol[143]. The movement of each tadpole was recorded on a digital camera (model EOS 70D, Canon USA Inc.) for 2 min on two occasions per day over 2 days. The videos were recorded at 30 frames per second with Full HD resolution (1920 × 1080 pixel^2) and using ALL-I compression. The percentage of time that each tadpole spent in total over the light and dark areas was calculated based on the location of the tadpoles' eyes by manually reviewing the recording, frame by frame, in Windows Media Player version 12 (Microsoft, Redmond, WA). A tadpole was considered to cross the white-black boundary only if both its eyes delineated the boundary.

Validation. The background color preference assay was validated using a positive control (+Ctrl.) group and by comparing results from our study to those independently established by others.[143] The control group (Ctr.) consisted of wild-type (WT) tadpoles (N = 16). The +Ctr group consisted of WT tadpoles (N = 4) one day after double axotomy of the left and right optic nerves following the protocol established in Ref. [143]. The movement of each tadpole was tracked during the background preference assay, as described earlier. Without visual perception, tadpoles were anticipated to randomly explore the white and black areas, viz. spending ~50% of time over either background color. Indeed, in an independent study, tadpoles post double axotomy spent 49–51% of time over the white background (Days 1 to 2), whereas the control explored this area in 99%–79% of time with this difference being statistically significant.[143] In our experiments, (**Fig. 3.2**), the mean time tadpoles spent over the preferred white background was ~56% in the positive control group vs. ~91% in the Ctrl. group with this difference amounting to statistical significance in our measurements (Mann-Whitney test, $p = 1.64\times10^{-5}$). Representative videos are provided for a Ctrl. and a +Ctrl. Tadpole (see SI in paper[26]). These results validated the sensitivity and robustness of the behavioral assay to support our studies on tadpole behavior.

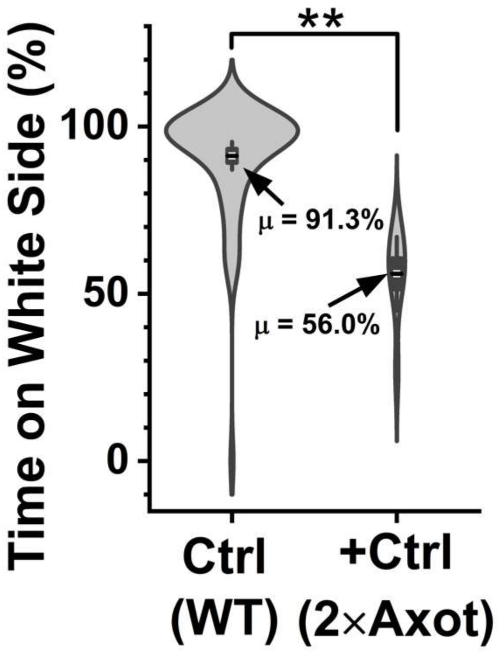


Figure 3.2 Validation of the behavioral assay. Double axotomy of the optic nerves significantly impacted the background color preference of the tadpoles. Key: *$p < 0.05$ (Mann-Whitney test).

The behaviors of experimental and control tadpoles were compared. The biological replicate size for the behavioral assay was N = 16 in the Ctr. group (tadpoles raised from nonmicrosampled embryos) and N = 15 in the experimental group (Exp., tadpoles raised from microsampled embryos). **Figure 3.3** presents representative images of the tadpoles. Power analysis was performed to compute the power to distinguish the wild-type group from the experimental group. Using two sample t-test ($\alpha = 0.05$, sample size = 60), the power was 0.08, which was below the acceptable power required to distinguish the two groups (> 0.8). Together with the Mann-Whitney test, these results indicated that microsampling imposed no detectable impact of statistical significance on the background preference of the tadpoles within the tested sensitivity of the behavioral assay.

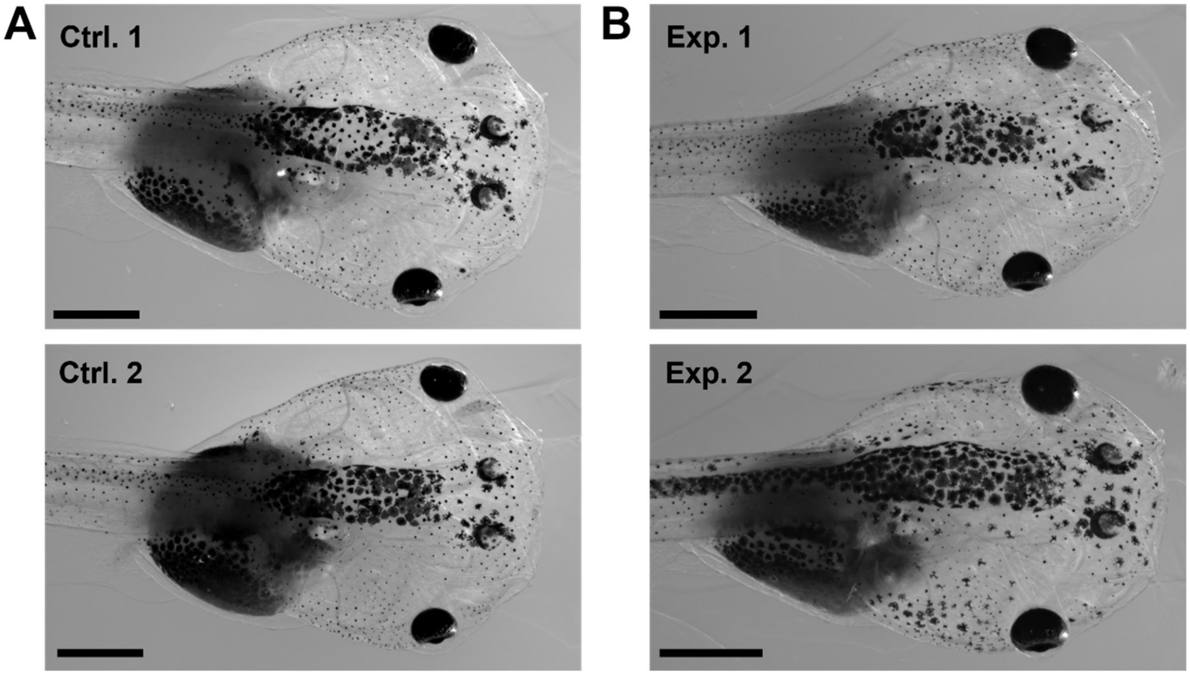


Figure 3.3 Representative high-resolution images of live tadpoles. Close-ups of tadpoles from **(A)** the control (Ctrl.) and **(B)** the experimental (Exp.) group. Scale bars = 1 mm.

3.3.6 Metabolomic measurement via CE-ESI-HRMS

Each subcellular metabolite extract was defrosted, and a 10 nL portion was analyzed on a custom-built CE-µESI platform with a high-resolution time-of-flight tandem mass spectrometer. This platform was built and operated according to our established protocols.[6, 49] The instrument was daily validated for the following analytical figures of merit before analyzing the cell extracts: lower limit of detection, 60 amol (acetylcholine); linear dynamic range of quantification, 4–5 log order (acetylcholine, methionine, and threonine); reproducibility, <1% relative standard deviation (RSD) with nonlinear time alignment[41] in separation time and <25% RSD in under-the-curve peak area. The experimental settings specific to this study were: CE, 100 cm capillary length and +22.5 kV potential (applied to the capillary inlet); ESI, coaxial sheath-flow CE-ESI interface (blunt-tip metal emitter, 1 µL/min sheath liquid, electrospray at –1,700 V, cone-jet spraying regime); MS, quadrupole orthogonal acceleration time-of-flight mass spectrometer with a collision induced dissociation cell (Impact HD, Bruker Daltonics, Billerica, MS) delivering 40,000 full width at half maximum (FWHM) resolution and <5 mDa *m/z* accuracy (both single- and tandem stage) using nitrogen gas at 12–18 eV collision energy. Each sample was analyzed in technical duplicate-triplicate.

3.3.7 Proteomic measurement via CE-ESI-HRMS

Each subcellular protein digest was defrosted and individually reconstituted in 2 µL of aqueous 60% acetonitrile containing 0.05% acetic acid. A 10 nL portion of each protein digest was analyzed on a custom-built CE-nanoESI platform equipped with a quadrupole orbitrap high-resolution mass spectrometer. The instrument was built and operated following our established protocols.[30, 51, 97] The platform was daily validated to deliver the following analytical figures of merit before measuring the cell extracts: lower limit of detection, 700 zmol (angiotensin II);

linear dynamic range of quantification, 4–5 log-order; reproducibility, <10% RSD in migration time and <25% RSD in signal intensity. The experimental settings specific to this study were: CE, 90 cm capillary length and +18 kV potential for 15 min and +15 kV thereafter (applied to the capillary inlet); CE-nanoESI[159], electrokinetic pump supplying 10% MeOH (0.05% AcOH) at +800–1,000 V (applied to the sheath reservoir), cone-jet spraying regime confirmed by a long working distance camera (EO-2018C with Mitutoyo Plan Apo objective, Edmund Optics, Barrington, NJ); and MS, quadrupole orbitrap mass spectrometer (Q-Exactive Plus, Thermo Scientific, Milford, MA).

Experimental conditions of data-dependent acquisition were tailored to fast and efficient separation by CE. Survey spectra were recorded at 35,000 FWHM (*m/z* 200) between *m/z* 350–1,600 (C-trap maximum injection time, 50 ms; AGC target, 1×10^6 ions). MS^2 scans were triggered for ion signal intensity exceeding 1.5×10^3 counts when the charge state was between +2–+7 and exhibited peptide-like isotopic distribution. An exclusion list of ~250 vitellogenin (Vtg) peptides was created based on identifications from a whole embryo "pre-run" to help minimize/avoid the sequencing of abundant yolk-related peptide signals. The MS^1 exclusion mass tolerance was 3 ppm. The MS^2 scans parameters were as follows: mass resolution, 17,500 FHWM (*m/z* 200); C-trap maximum injection time, 60 ms with AGC target, 5×10^4 counts; dynamic exclusion time and mass tolerance, 9 s and 5 ppm; peptide isolation window, 1 Da; collision parameters, HCD in nitrogen at 28% normalized collision energy (NCE). Each protein digest was measured in technical duplicate-triplicate.

To increase identification numbers using the match between run feature, we also measured cell samples using our CE-nanoESI coupled to a Fusion Lumos quadrupole-linear ion trap-orbitrap tribrid mass spectrometer (Thermo Scientific). MS^1 parameters were as follows: mass analyzer, orbitrap; mass range, *m/z* 350–1,600; mass resolution; 50,000 FWHM (*m/z* 200); C-trap maximum injection time, 86 ms; automatic gain control (AGC) target, 1×10^6 ions. MS^2 scans were triggered for ion signal intensity exceeding 1.5×10^3 counts, with charge state between 2–7, and exhibited peptide-like isotopic distribution. The MS^2 scans parameters were as follows: mass analyzer, ion trap; AGC target, 5×10^4 counts; dynamic exclusion time and mass tolerance, 9 s and 5 ppm; peptide isolation window, 1 Da; collision parameters, HCD in nitrogen at 32% normalized collision energy.

3.3.8 Identification and quantification of metabolites and proteins

Metabolites were identified and relatively quantified following our established protocols.[6, 43, 49] Molecular features (signals with unique *m/z* vs. migration time values) with $S/N \geq 3$ were surveyed between *m/z* 50–500 using a semi-automated approach. For each molecular feature, the under-the-curve peak area was integrated, providing a proxy for concentration. These quantitative metadata were imported to MetaboAnalyst version 5.0[160] for mean-normalization and log-transformation, then exported for integration with the proteomics metadata, before importing of the unified metabolomic-proteomic metadata for statistical data analysis. Selected molecular features were identified by matching the accurate mass, isotope distribution, CID fingerprint, and migration time of the unknown signal to data available in metabolomics MS-MS/MS databases (Metlin[161]) or experimentally determined by us in this study or previously[43, 50].

Proteins were identified and quantified following established protocols. The MS-MS/MS data were matched against the *X. laevis* proteome using MaxQuant[16] version 1.6 running the Andromeda search engine.[15] The *X. laevis* proteome database was custom built by concatenating the mRNA-derived PHROG1r0 database (downloaded from reference[162]) and the *X. laevis* SwissProt proteome database (downloaded from UniProt on Oct. 2019). The following search parameters were used: no fixed modification; variable modifications, methionine oxidation, and asparagine and glutamine deamidation; minimum peptide length, 5 amino acids; MS^1 mass deviation, 5 ppm; MS^2 mass tolerance, 20 ppm for first search and 10 ppm for de novo tolerance; search for common contaminants enabled; match between run enabled with 5 min time shift tolerance. A minimum of one unique peptide was required for successful protein identification. Identified common contaminants were manually excluded from proteins identified or quantified in this study. Peptide and protein identifications were filtered to <1% false discovery rate (FDR) calculated against a reversed-sequence decoy database. Proteins were quantified using the label-free quantification (LFQ) approach in MaxQuant using unique and razor peptides.[163] Only proteins with less than 50% missing LFQ values across all the samples were considered as successfully quantified and used for statistical and pathway analysis. Protein LFQ values were mean-normalized and log-transformed in MetaboAnalyst version 5.0[160], then exported for integration with the metabolomics data, before importing of the unified metadata for statistical analysis.

3.3.9 Experimental design and statistics

For each portion of this study, the number of necessary technical and biological replicates for statistical significance were determined based on a pilot study and the analytical performance metrics of the custom-built CE-ESI-MS platform. Statistical, multivariate, and joint pathway analyses were performed in MetaboAnalyst 5.0. Statistical significance was marked at $p \leq 0.05$ (paired student t-test for normally distributed data). Power analysis (two-sample t-test, alpha value 0.05, hypothetical sample size N = 60), Student's t-test (parametric, normally distributed data), Mann-Whiney test (nonparametric, non-paired data), and Wilcoxon signed rank test (nonparametric, paired data) were performed in OriginPro 2020b (OriginLab Corp., Northampton, MA). For pathway enrichment analysis, gene and metabolite centered pathway was chosen. Enrichment was measured using hypergeometric test, topology was calculated by degree of centrality, and the integration method used was combined *p*-value (overall).

3.4 Results and discussion

We developed the bioanalytical technology and demonstrated its use in molecular systems cell biology with compatibility for cell-, neurodevelopmental, and behavioral biology. **Figure 3.4** presents our analytical and biological tracks, essentially connecting the physical and life sciences. To enable molecular network analyses in live embryos, proteins and metabolites were measured by balancing sensitivity and developmental viability. In the context of neurodevelopment, the analysis is considered *in vivo* if the cell is able to heal its membrane, divide, and form progenies for tissue specification and development to give way to a normally behaving organism, viz. a tadpole. The biological trace in the study determined this biological phenotype on the basis of morphological, survival, and behavioral assays. By design, the integration of these tracks generates multi-dimensional metadata to open a window into molecular systems biology and help develop new knowledge and support hypothesis-driven studies.

3.4.1 In vivo microsampling

The vertebrate South African clawed frog (*X. laevis)* was an ideal biological model for technology development, validation, and application in this study. Each cell in the cleavage-stage frog embryo is identifiable based on pigmentation and location and has reproducible tissue fates.[164] **Figure 3.4** exemplifies the identification of the left dorsal-animal (L-D1) and left ventral-animal (L-V1) cells in a live 8-cell embryo, which respectively form the neural and epidermal tissues. *X. laevis* embryos develop externally to the mother and contain considerably large cells (180 nL volume/cell in the 8-cell embryo), permitting direct inspection and access to the cells. However, the cells populate the embryo in a complex three-dimensional morphology (spherical arrangement) and divide rapidly, in ~15–25 min per cell cycle between the 2- and 32-cell stages.[139] These physical and characteristics combined with the transient nature of an "embryonic body" fundamentally challenge existing single-cell HRMS technologies in terms of scalability for spatial, temporal, and *in vivo* operation.

We recognized spatiotemporally scalable sampling as indispensable for this study. While dissociation[79] and manual dissection[6, 30, 100] permit cell isolation from the embryo, the former loses spatial information on cell heterogeneity, and the latter requires substantial dexterity and lacks scalability; these features both have important implications in cell and developmental biology. With analyses occurring *ex vivo*, the current tools exclude the possibility of biological studies over critical developmental stages, ranging from gastrulation and neurulation to organogenesis and metamorphosis. Likewise, these strategies are incompatible with assessing function and behavior at an organismal level, such as the sentient tadpole. Currently, there exists no single-cell HRMS technology capable of determining the proteo-metabolome of single cells embedded in complex tissues or organisms *in vivo* with scalability.

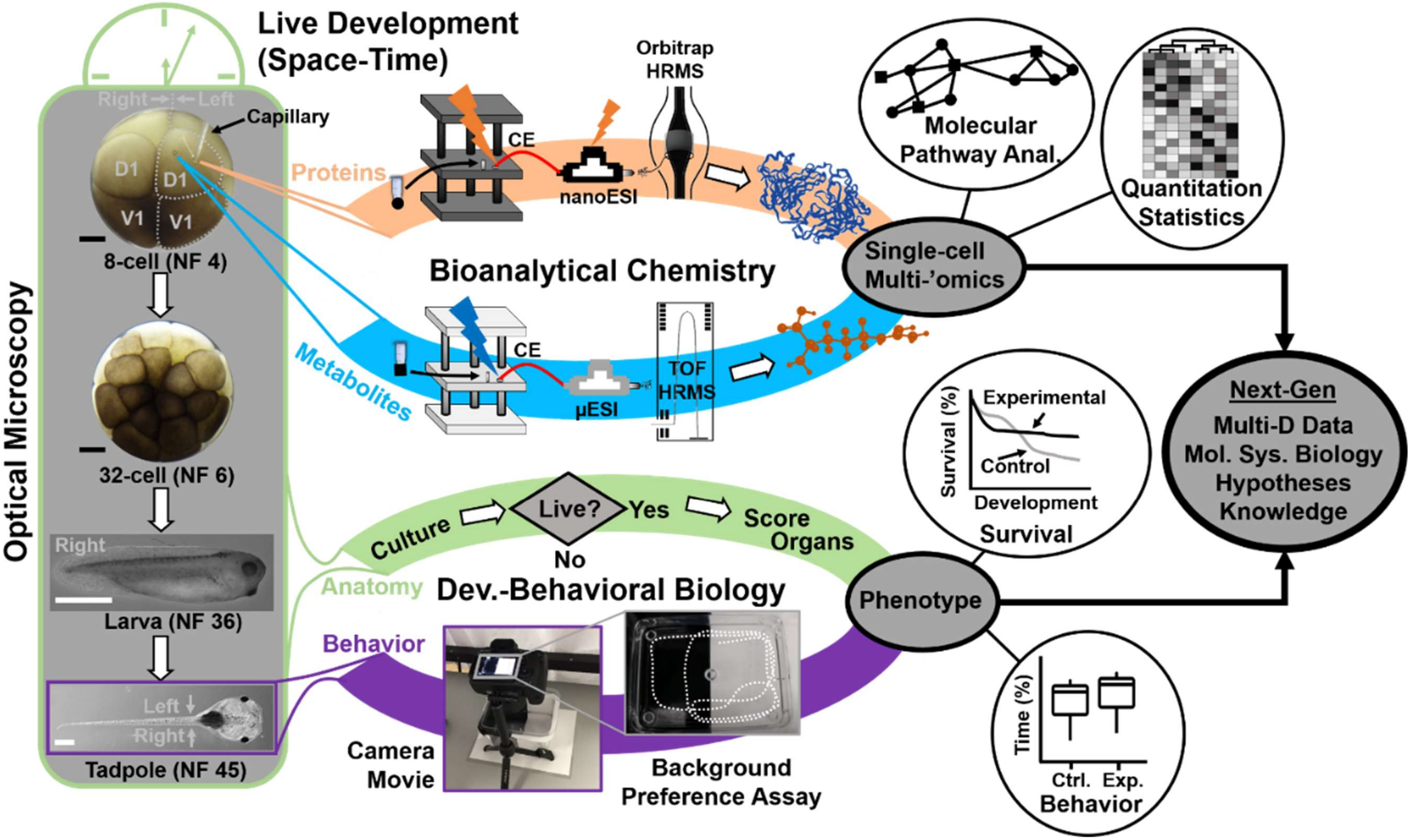


Figure 3.4 Interdisciplinary strategy enabling *in vivo* subcellular proteo-metabolomic systems biology with demonstrated compatibility for cell-, neurodevelopmental, and behavioral biology using *X. laevis* as the biological model. As an example, the left dorsal-animal (L-D1) and left ventral-animal (L-V1) cells were identified and the content of each cell rapidly microsampled twice using clean capillaries each time. The collected protein and metabolite samples were analyzed using custom-built ultrasensitive capillary electrophoresis (CE) nano/micro-flow electrospray ionization (ESI) high-resolution orbitrap and time-of-flight (TOF) mass spectrometry (HRMS). The tadpoles developing from the embryos were characterized for survival, anatomy, and behavior. Scale bars = 250 µm (black), 1 mm (white).

We proposed double capillary microsampling as a potential solution for analyzing the cells *in vivo*. To serve as a micropipette for sample collection, a borosilicate capillary (0.75/1 mm inner/outer diameter) was flame-pulled, then cleaved to a 20-µm tip diameter (see **3.2 Experimental Section**), and connected to a microinjector delivering negative pressure pulses. The tip of the capillary was inserted into the cell of interest under translation by a precision micromanipulator and guidance by real-time stereomicroscopy. In each of N = 20 stereotypical 8-cell embryos (experimental group), the L-D1 and L-V1 cells were each microsampled twice in a random order with both cells from the same embryo and one sampling for each 'omic analysis. By performing the single-cell analyses exclusively on one side of the embryo (the left, experimental), tissues and organs arising from the nonsampled cells on the other side (right D1, R-D1 and right V1, R-V1) served as the internal control in each embryo to facilitate analysis of tadpole anatomy (see later). For the control group (Ctrl.), N = 20 stereotypical 8-cell embryos were cultured under identical conditions, without microaspiration. This approach was inspired by *in situ* microprobe sampling that we recently developed for cells in zebrafish and *X. laevis* embryos, albeit functioning only *ex vivo* and exclusively for metabolomics[43, 49, 121] or proteomics[52], but not both 'omes at the same time or on the same cell. We experimentally tested that double microsampling using larger capillaries (e.g., ~80 µm tips) and/or longer aspiration times (scalability) allowed us to aspirate ~100 nL, viz. 50% of the cell volume. These amounts were sufficient for single-cell HRMS; however, the embryos failed to survive the analysis due to substantial damage by the large capillary (see **Fig. 3.1**).

3.4.2 Survival, morphological, and behavioral assays

To enable cell sampling *in vivo* in the live embryo, we scaled the approach with assistance from survival assays. **Figure 3.1** presents significant membrane damage on the microsampled cells in a 16-cell embryo before refinement of this approach. In a series of experiments (data not shown), the tip size of the microprobe, the volume collected, and the speed of microsampling were tailored to the cells while monitoring the success of cell divisions after sampling under the stereomicroscope. Each sampling employed a clean and unused capillary to aid cell and embryonic survival by avoiding accidental contamination. **Figure 3.5A** tracks the percentage of embryos surviving after ~10 nL, or ~5% of the cell volume, was microaspirated twice from the L-D1 and L-V1 cells (N = 20 embryos). Double microsampling took less than 30 s per cell. By using 20-μm-tip capillaries and conducting two independent sampling events in each cell instead of a single one, we were able to promote cell viability and microhandling of viscous cell contents for downstream sample processing. Embryonic survival was quantified over 6 key stages of vertebrate development: cleavage (Nieuwkoop-Faber, NF 6), gastrula (NF 10.5, **Fig. 3.5B**), neurula (NF 13, **Fig. 3.5B**), early tailbud (NF 22), late tailbud (NF 36 and 41), and tadpole (NF 45, **Fig. 3.5B**). Compared to 100% of tadpoles successfully developing in the control group (unperturbed wild-type, WT) at all these stages of development, 100% of the embryos developed to the neurula stage and 95% survived to tadpoles in the experimental group. Therefore, capillary microsampling was successfully scaled to preserve viability.

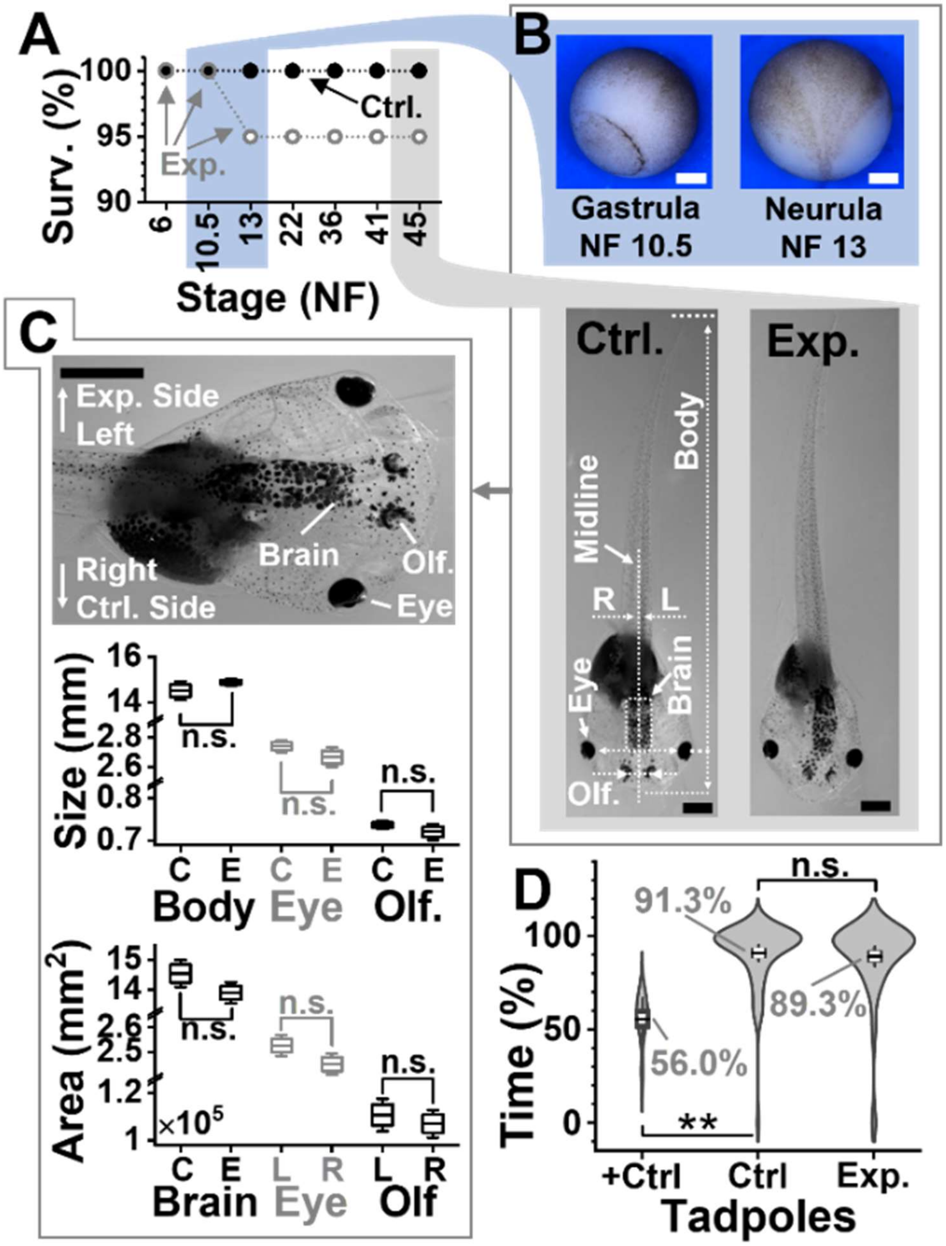


Figure 3.5 Developmental and behavioral impact of the technology. **(A)** Survival analysis (Surv.) at key developmental stages (NF, Nieuwkoop-Faber) for $N = 20$ embryos in the control (Ctrl.) group and $N = 20$ embryos in the experimental. **(B)** Representative (**top panel**) Ctr. gastrulae and Ctrl. neurula and (**bottom**) Ctr. and experimental (Exp.) tadpoles. Shown: midline separating left (L)-right (R) axis, body length, brain area, and distance between eyes and olfactory (Olf.) organs. **(C)** Comparison of tadpole anatomy. (**Top**) Close-up image of a tadpole showing the experimental (left) and control (right) side. Right tissues are labeled. (**Middle**) Analysis of organ size on the left (experimental side) between $N = 5$ Ctrl. (C) and $N = 6$ Exp. (E) surviving tadpoles randomly selected. Key: n.s., non-significant, *$p < 0.05$, Mann-Whitney test. (**Bottom**) Analysis of tissue areas for total brain (Ctr. vs. Exp. group), the eyes and, the olfactory organs in the 6 Exp. tadpoles. Key: n.s., non-significant, Wilcoxon signed rank, paired). **(D)** Comparison of visual function in a background-color preference assay for $N = 16$ Ctrl. vs. $N = 15$ Exp. tadpoles. Assay validation via double axotomy of the optic nerves in $N = 4$ positive Ctrl. tadpoles (+Ctrl.). Key: n.s., non-significant; *$p < 0.05$, **$p < 0.005$ (Mann-Whitney). Scale bars, 250 µm (white) and 1 mm (black). Box-whisker plots: Box, 1×standard error of the mean (SEM), whiskers: 1.5×SEM.

Anatomy was also analyzed for the surviving tadpoles. **Figure 3.5B** presents typical examples for tadpoles from the control and experimental groups. Their morphologies were characterized in terms of whole-body length as well as the sizes of the eyes, olfactory organs (Olf.), and total brain, which partially derive from the L-D1 and L-V1 cells (**Fig. 3.5B**). Body symmetry was also analyzed for the eyes and olfactory tissues along the midline separating the left (experimental) side and the right (control) side (see close-up in **Fig. 3.5C**). As shown in **Figure 3.5C**, the body length ($p = 0.315$, Mann-Whitney) as well as the center-to-center distances between the eyes ($p = 0.411$, Mann-Whitney) and the olfactory tissues ($p = 0.523$, Mann-Whitney) were indistinguishable, as was the total area of the brain ($p = 0.121$, Mann-Whitney). The size (area) of the eyes ($p = 0.142$, Wilcoxon signed rank, paired) and olfactory organs ($p = 0.142$, Wilcoxon signed rank, paired) were also indistinguishable between the experimental and control sides in the tadpoles. Therefore, double subcellular capillary microsampling of the L-D1 and L-V1 in the cleavage-stage embryo led to no detectable impairment on tadpole development, morphologically.

Physical appearance, however, does not equate to performance; therefore, we further evaluated the animals based on behavior. The visual (also including motor, sensory, and neural processing) function was compared at stages 45–50, when tadpoles display a robust preference for lighter background in a color preference assay[143] (see setup in **Fig. 3.4**). Indeed, as shown in **Figure 3.5D**, when presented with a dark vs. light background in a tank, the control (WT) tadpoles resided ~91% of the time over the light background (N = 16). These tadpoles recapitulated this robust behavioral phenotype at this stage,[143] thus confirming the robustness of the assay in our hands. We further validated the sensitivity of the assay to detect visual impairment using a positive control (+Ctrl). Tadpoles that underwent double axotomy of their optic nerves (N = 4) explored the light background only at ~56% of the time, in a significant difference from the control group ($p = 1.64\times10^{-5}$). Indeed, without vision, tadpoles are anticipated to make random explorations, spending ~50% on either side of the tank. These results also agreed with results from an independent investigation[143], thus confirming that the assay was sufficiently sensitive to detect visual function in our experiments. Swimming ~89% time over the light area, the experimental group (N = 15) presented indistinguishable visual behavior from the control tadpoles ($p = 0.41$, Mann Whitney). Double microsampling of each of the 2 cells in the same embryo preserved embryonic viability and organismal behavior. The approach opened, for the first time in single-cell HRMS, the possibility of dual 'omics of single cells in a live organism, which also undergoes dynamic spatial and temporal reorganization to form a behaving animal, as demonstrated here for the *X. laevis* embryo and its tadpole *in vivo* in this work.

3.4.3 Proteomic and metabolomic analysis of collected cellular contents

Detection of the limited amounts of proteins and metabolites from the cells necessitated ultrasensitive analysis. From each cell type in the same embryo (N = 4), the double aspirates were independently processed via a proteomic and metabolomic workflow (recall **Fig. 3.4**). To eliminate potential systematic biases during sample collection, we randomized the location of subcellular sampling and the order of aspirating the protein and metabolite samples. Each embryo, cell, and processed protein and metabolite sample was assigned a unique identifier in our study, although this information was purposefully not used during data analysis to eliminate potential (conscious or subconscious) biases during sample preparation, processing, or instrumental measurement. This information was only revealed after data collection to aid results interpretation. As detailed in the **3.2 Experimental Section**, each subcellular sample was processed to yield metabolic and proteomic microextracts. For bottom-up proteomics, potential protein losses on vial and pipette tip surfaces were minimized by eliminating the traditional steps of desalting, alkylation, and reduction. The final subcellular samples contained the tryptic digest in 2 µL for bottom-up proteomics and 4 µL for metabolomics. To avoid systematic bias during instrumental measurement, the subcellular ʻomes were blinded for cell type and analyzed in a randomized order.

With ~100–10,000-times smaller amounts of material collected from the cells in this study than those typically processed/measured by standard HPLC-HRMS, we turned to our ultrasensitive CE-ESI-HRMS instruments for assistance. These platforms provided ~700 zmol sensitivity for peptides on a tribrid quadrupole-orbitrap-ion trap[30, 52] and ~60 amol sensitivity for metabolites on a quadrupole time-of-flight high-resolution[6, 43] mass spectrometer. To facilitate the instrumental measurements, the metabolomic analytical track was limited to cationic electrophoresis in this study. Dual cationic-anionic metabolomics is also possible on these instruments to deepen the detectable portion of the single-cell metabolome.[50] Measurement of ~0.5% (v/v) of the protein extract, thus ~0.03% of the cell's volume, identified 738 proteins (false discovery rate < 1%). Similarly, analysis of ~0.25% (v/v) of the metabolite sample (volume), viz. ~0.01% of the cell's volume, produced ~150 nonredundant metabolic molecular features, 57 of which were identified with a high Level-1[12] confidence, listed in **Table 3.1** Notably, detection of these ~800 proteins and metabolites was achieved in an untargeted (discovery) manner, requiring no prior knowledge of cellular composition or the use of functional probes; antibodies or antisense oligos were unnecessary for our technology.

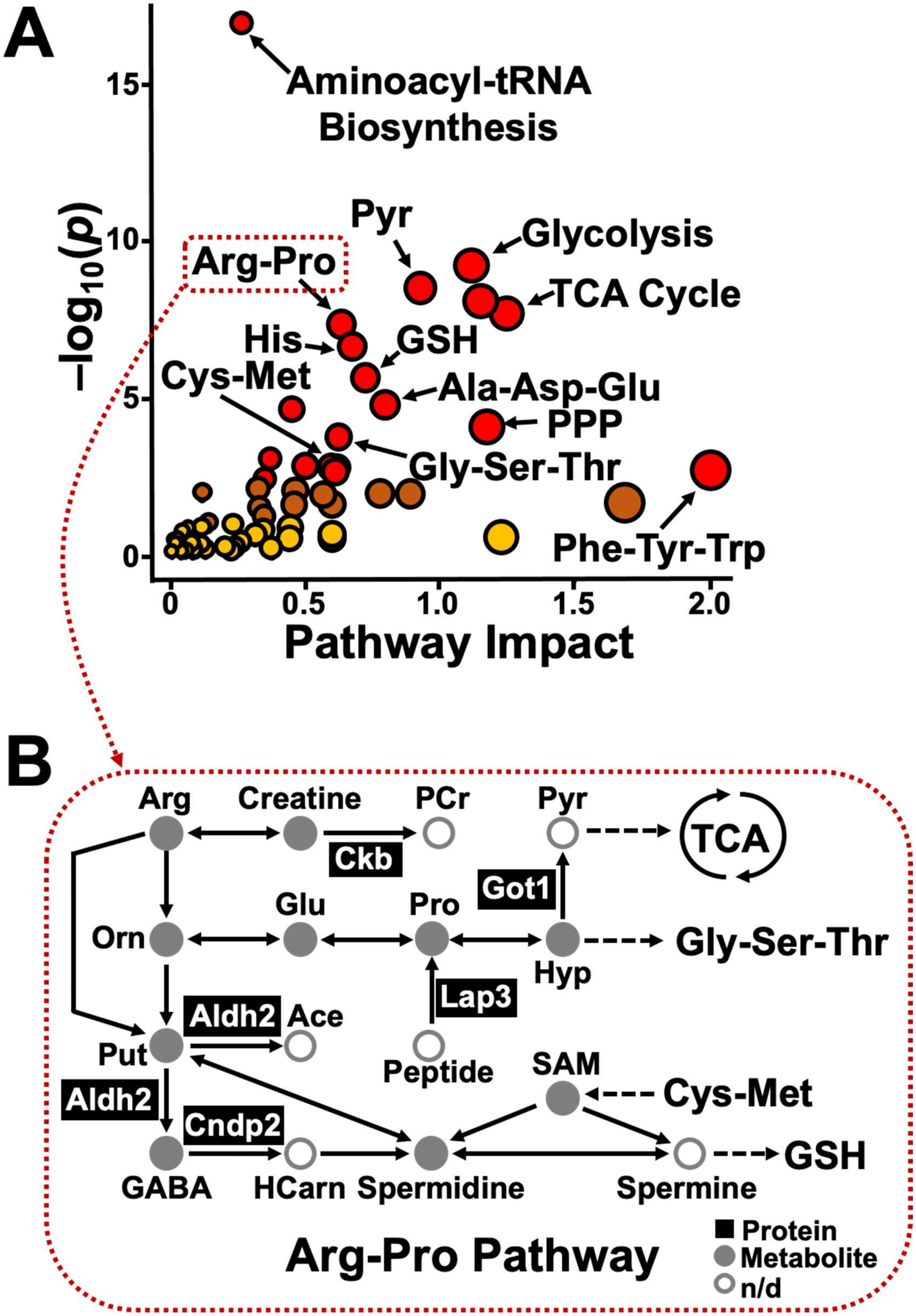


Figure 3.6 Qualitative and quantitative proteo-metabolomic systems biology on the L-D1 and L-V1 cell in the live embryo. **(A)** KEGG pathway analysis of the identified proteins and metabolites. Pathway data are in **Table 3.2**. **(B)** Example showing proteins and metabolites with complementary enrichment in the arginine-proline pathway.

3.4.4 Pathway and statistical analyses using proteo-metabolomic datasets

These multi-omics data also opened a window into (sub)cellular biochemistry, directly in the live embryo. Statistical enrichment analysis using the Kyoto Encyclopedia of Genes and Genomes[165] revealed significant coverage of several known pathways (**Fig. 3.6A**). For example, the arginine-proline pathway, TCA cycle, glycolysis, and pyruvate metabolism were represented with statistical significance, whereas several other amino acid pathways had high pathway impact. Notably, these proteins and metabolites assessed cellular biochemistry in complementary performance. For example, the biosynthesis of aminoacyl-tRNA and valine-leucine-isoleucine was only represented by metabolites, whereas glycolysis/gluconeogenesis was solely enriched for by proteins (see statistics and fold changes in **Table 3.2**. Further, other pathways, such as arginine biosynthesis and arginine-proline metabolism, were represented by both 'omes (**Table 3.2**. **Figure 3.6B** annotates arginine-proline metabolism based on the complementary detection of proteins and metabolites in the cells. The observed coverage of the network raises a potential for studies targeting particular 'omic processes. By also including anionic metabolites during CE-ESI-HRMS in future experiments, we anticipate many other cellular pathways to be detectable by this approach, such as the TCA cycle and energy production as well as drug metabolism. These results illustrate the benefit of measuring more than one 'ome in single cells toward a holistic understanding of cellular systems biology at the molecular level.

Finally, we applied the technology to ask whether the pluripotent cells harbored detectable differences in proteo-metabolomic activity at such an early stage of development. Although embryonic cells have reproducible tissue fates in the cleavage-stage embryo of *Xenopus*, unbiased and parallel characterization of their proteome and metabolome has not been possible due to a lack of sufficiently sensitive and specific molecular tools. We relatively quantified ~450 proteins and ~80 metabolites between the L-D1 and L-V1 cells in each embryo (N = 4) based on the subcellular HRMS data obtained in this study. Quantification was based on signal abundances in the HRMS datasets, which offer an effective and robust proxy for concentration in CE-ESI-HRMS.[30, 41] Quantified protein and metabolite signals are presented in **Appendix 3.1**. As highlighted earlier, the phenotypes of the cells or the identifier of each embryo were hidden to ensure bias-free data analysis. **Figure 3.7** presents the unsupervised hierarchical cluster analysis (HCA) and intensity heat map of the data based on the top 40 statistically most significantly dysregulated molecules, including 13 metabolites and 27 proteins.

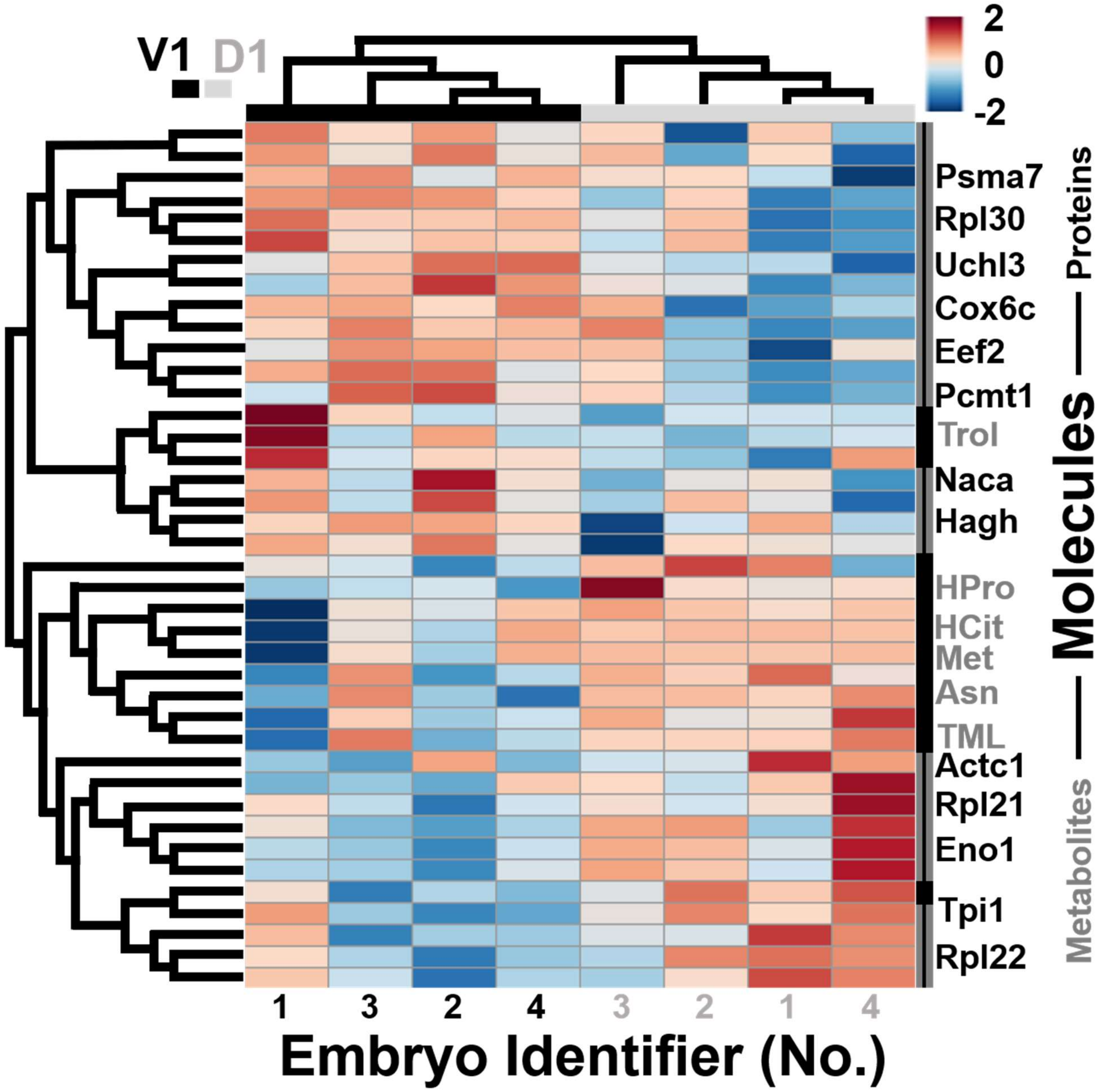


Figure 3.7 Unsupervised (blinded) hierarchical cluster–heat map analysis of the detected protein and metabolite quantities revealing reproducible, systematic, and significant differences between the cell types. Top 40 most significantly dysregulated molecules shown ($p < 0.05$, see quantitative statistics in **Appendix 3.1**).

The pathway data are listed in **Table 3.2**. The top dendrogram clusters the samples into two groups in the HCA plot. Upon revealing the identity of the samples, we learned that the groups corresponded to the D1 and V1 cell types. Further, the heatmap revealed notable quantitative molecular differences between the cell types. The left dendrogram organized the proteins and metabolites into two groups in the HCA plot. The HCA heatmap readily captures reproducible abundance differences between the cell types. Specifically, in this map, 20 molecules, containing 10 metabolites and 10 proteins, were more abundant in the L-D1, whereas 20 molecules, encompassing 3 metabolites and 17 proteins were enriched in the L-V1. **Appendix 3.1** tabulates the proteins and metabolites with statistically ($p \leq 0.05$) and biologically significant dysregulation ($|\text{fold change}| \geq 1.5$). These intriguing molecular differences would have been lost due to averaging across a large number of cells pooled for classical HRMS approaches.

3.5 Conclusions

We enabled *in vivo* single-cell proteomics and metabolomics in the same single cell in live chordate embryos using *X. laevis* as the biological model. The technology revealed quantitative proteo-metabolomic differences between cells at the cleavage stage. These results afford previously unavailable insights into the establishment of cell heterogeneity during embryogenesis; they also challenge our current understanding of the underlying molecular processes. Differential production of these molecules reveals that asymmetry along the dorsal-ventral, one of the three primary axes of the vertebrate body plan, is set up rather early during embryonic development, when transcriptional heterogeneity is not detectable along this axis based on sequencing of the respective single-cell transcriptomes[69, 166] in *Xenopus*. These findings support our earlier discovery of molecules, such as metabolites[6], affecting tissue fates via noncanonical mechanisms of action. Although determining the exact biological significance of these findings goes beyond the scope of this study, the data generated in this work may be used to develop hypotheses for experiments targeting specific proteins and metabolites and their functions.

Supporting future investigations, our technology is scalable in space and time to other types of cells and different biological models. Subcellular HRMS is compatible with complex tissues and live development, as demonstrated for the live embryo here. It does not escape our attention that our technology can be used to perform multi-‘omics on subcellular organelles. Further, the approach is also adaptable to classical and modern tools of biology and health studies, such as optical microscopy and behavioral assays (as demonstrated in this study) to characterize phenotypes as well as established or contemporary tools of molecular biology, including expression/translation-blocking morpholinos or gene editing by CRISPR-Cas9, to probe function. We anticipate *in vivo* proteo-metabolomic subcellular CE-ESI-HRMS to expand the contemporary toolbox of cell and developmental biology, neuroscience, and health research. This technology realizes new strengths in the molecular study of the building block of life, the cell, and processes governing the formation of tissues, organs, and entire organisms with complex behavior.

Table 3.1 Metabolites identified in single L-D1 and L-V1 cells in 8-cell *X. laevis* embryos using *in vivo* subcellular HRMS.

ID	**Metabolite Name**	**Abbreviation**	**Formula**	**t_m**	***m/z***	***m/z***	**Δ *m/z***	
				(min)	**Measured**	**Theoretical**	**(mDa)**	**(ppm)**
1	Spermidine[**, †]		$C_7H_{19}N_3$	8.8	146.1651	146.1652	0.1	0.7
2	Putrescine[‡]		$C_4H_{12}N_2$	9.3	89.1070	89.1073	0.3	3.4
3	Methylhistamine[**, †]		$C_6H_{11}N_3$	9.8	126.1019	126.1026	0.7	5.6
4	Thiamine[**, †]		$C_{12}H_{16}N_4OS$	13.6	265.1111	265.1118	0.7	2.6
5	Choline[*,**, †]	Cho	$C_5H_{13}NO$	14.1	104.1066	104.1070	0.4	3.8
6	Ser-Arg[**, †]		$C_9H_{19}N_5O_4$	14.9	262.1513	262.1510	–0.3	–1.1
7	S-adenosyl-methionine[**, †]	SAM	$C_{15}H_{22}N_6O_5S$	15.0	399.1457	399.1445	–1.2	–3.0
8	Ornithine[**, †]	Orn	$C_5H_{12}N_2O_2$	15.1	133.0983	133.0972	–1.1	–8.3
9	Lysine[*,**, †]	Lys	$C_6H_{14}N_2O_2$	15.3	147.1129	147.1128	–0.1	–0.7
10	Sarcosine[‡]	Sar	$C_3H_7NO_2$	15.3	90.0542	90.0550	0.8	8.9
11	Arginine[*,**]	Arg	$C_6H_{14}N_4O_2$	15.8	175.1185	175.1190	0.5	2.9
12	Histidine[*,**, †]	His	$C_6H_9N_3O_2$	16.0	156.0768	156.0768	0.0	0.0
13	N6, N6, N6-trimethyl-lysine[**, †]	TML	$C_9H_{20}N_2O_2$	16.0	189.1592	189.1598	0.6	3.2
14	Homolysine[**, †]		$C_7H_{16}N_2O_2$	16.3	161.1306	161.1285	–2.1	–13.0
15	Methylhistidine[**, †]		$C_7H_{11}N_3O_2$	16.5	170.0937	170.0924	–1.3	–7.6
16	Urocanate, cis-[**, †]	cURA	$C_6H_6N_2O_2$	16.8	139.0498	139.0502	0.4	2.9
17	Acetylcholine[*,**]	AcCho	$C_7H_{15}NO_2$	16.9	146.1173	146.1176	0.3	2.1
18	Trolamine[‡]	TEA	$C_6H_{15}NO_3$	17.1	150.1134	150.1125	–0.9	–6.0
19	γ-aminobutyrate[*,**, †]	GABA	$C_4H_9NO_2$	17.3	104.0707	104.0706	–0.1	–1.0
20	Urocanate, -trans[**, †]	tURA	$C_6H_6N_2O_2$	17.5	139.0497	139.0502	0.5	3.6
21	Guanine[**, †]		$C_5H_5N_5O$	17.5	152.0569	152.0567	–0.2	–1.3
22	Methylguanine[**, †]		$C_6H_7N_5O$	17.8	166.0714	166.0723	0.9	5.4
23	Carnitine[*,**, †]	Car	$C_7H_{15}NO_3$	17.9	162.1128	162.1125	–0.3	–1.9
24	Pyridoxal (Vitamin B6)[**, †]		$C_8H_9NO_3$	18.4	168.0654	168.0655	0.1	0.6

25	Acetylcarnitine[*,**, †]	AcCar	$C_9H_{17}NO_4$	19.3	204.1227	204.1230	0.3	1.5
26	Methylaspartate[**, †]		$C_5H_9NO_4$	19.9	148.0604	148.0604	0.0	0.0
27	Niacinamide[*, †]		$C_6H_6N_2O$	20.2	123.0559	123.0553	–0.6	–4.9
28	Creatine[*,**, †]	CR	$C_4H_9N_3O_2$	20.2	132.0767	132.0768	0.1	0.8
29	Alanine[*, †]	Ala	$C_3H_7NO_2$	20.2	90.0548	90.0550	0.2	2.2
30	Propionylcarnitine[**, †]		$C_{10}H_{19}NO_4$	20.6	218.1368	218.1387	1.9	8.7
31	Pro-Val[**, †]		$C_{10}H_{18}N_2O_3$	20.8	215.1383	215.1393	1.0	4.6
32	Adenosine[**]		$C_{10}H_{13}N_5O_4$	22.1	268.1031	268.1040	0.9	3.4
33	Argininosuccinate[**, †]		$C_{10}H_{18}N_4O_6$	22.7	291.1295	291.1299	0.4	1.4
34	Valine[*,**, †]	Val	$C_5H_{11}NO_2$	24.6	118.0855	118.0863	0.8	6.8
35	Serine[*,**, †]	Ser	$C_3H_7NO_3$	24.7	106.0493	106.0499	0.6	5.7
36	Isoleucine[*, †]	Ile	$C_6H_{13}NO_2$	24.9	132.1014	132.1019	0.5	3.8
37	Leucine[*, †]	Leu	$C_6H_{13}NO_2$	25.1	132.1014	132.1019	0.5	3.8
38	Asparagine[*,**, †]	Asn	$C_4H_8N_2O_3$	26.2	133.0612	133.0608	–0.4	–3.0
39	Threonine[*,**, †]	Thr	$C_4H_9NO_3$	26.2	120.0650	120.0655	0.5	4.2
40	Methionine[*,**]	Met	$C_5H_{11}NO_2S$	26.8	150.0594	150.0583	–1.1	–7.3
41	Glutamine[*,**, †]	Gln	$C_5H_{10}N_2O_3$	26.9	147.0766	147.0764	–0.2	–1.4
42	2-Aminoadipate[**, †]		$C_6H_{11}NO_4$	27.0	162.0758	162.0761	0.3	1.9
43	Citrulline[**, †]		$C_6H_{13}N_3O_3$	27.3	176.1025	176.1030	0.5	2.8
44	Glutamate[*,**, †]	Glu	$C_5H_9NO_4$	27.3	148.0609	148.0604	–0.5	–3.4
45	Homocitrulline[**, †]		$C_7H_{15}N_3O_3$	27.3	190.1178	190.1186	0.8	4.2
46	Acetyllysine[†]		$C_8H_{16}N_2O_3$	27.6	189.1255	189.1234	–2.1	–11.1
47	Phenylalanine[*,**, †]	Phe	$C_9H_{11}NO_2$	27.8	166.0865	166.0863	–0.2	–1.2
48	Tyrosine[*,**, †]	Tyr	$C_9H_{11}NO_3$	28.2	182.0806	182.0812	0.6	3.3
49	Hypoxanthine[**, †]	HPX	$C_5H_4N_4O$	28.6	137.0454	137.0458	0.4	2.9
50	Proline[*,**, †]	Pro	$C_5H_9NO_2$	28.6	116.0703	116.0706	0.3	2.6
51	Ser-Val[‡,†]		$C_8H_{16}N_2O_4$	29.0	205.1181	205.1183	0.2	1.0
52	Cysteine[*,**, †]	Cys	$C_3H_7NO_2S$	30.0	122.0269	122.0270	0.1	0.8

53	Aspartate*,**,†	Asp	$C_4H_7NO_4$	30.2	134.0445	134.0448	0.3	2.2
54	Glycine betaine*		$C_5H_{11}NO_2$	30.8	118.0856	118.0863	0.7	5.9
55	Glutathione, oxidized**	GSSG	$C_{20}H_{32}N_6O_{12}S_2$	31.4	307.0838	307.0833	–0.5	–1.6
56	Hydroxyproline**,†		$C_5H_9NO_3$	32.6	132.0652	132.0655	0.3	2.3
57	Glutathione, reduced*,**	GSH	$C_{10}H_{17}N_3O_6S$	34.4	308.0909	308.0911	0.2	0.6

Notes: The exact mass was calculated using IsotopePattern (Bruker Daltonics, Billerica, MA). Asterisk (*) denotes identification was based on migration time comparison to chemical standards. Double asterisk (**) marks identifications supplemented with tandem MS experiments on chemical standards and/or comparison with MS/MS data available at METLIN. Dagger (†) means metabolite was identified in references[43, 50]. Double dagger (‡) means identification was based on accurate mass matching with METLIN database (5 ppm accuracy).

Table 3.2 The top 10 most significant pathways enriched for the metabolites or proteins detected in the L-D1 and L-V1 cells.

Pathway	Expected Total	Hit Metabolites	Hit Proteins	*P*-value	FDR	Impact
Aminoacyl-tRNA biosynthesis	74	18	0	9.91E-18	6.44E-16	0.26
Glycolysis or Gluconeogenesis	61	0	18	1.09E-09	3.53E-08	1.12
Pyruvate metabolism	45	0	14	5.12E-09	1.11E-07	0.93
Arginine biosynthesis	27	7	4	1.45E-08	2.35E-07	1.15
Citrate cycle (TCA cycle)	42	0	13	3.22E-08	4.19E-07	1.24
Arginine and proline metabolism	78	9	9	7.79E-08	8.43E-07	0.64
Histidine metabolism	32	6	5	3.60E-07	3.34E-06	0.68
Glutathione metabolism	56	6	8	3.71E-06	3.01E-05	0.73
Alanine, aspartate, and glutamate metabolism	61	7	5	2.86E-05	2.06E-04	0.80
Valine, leucine, and isoleucine biosynthesis	12	4	0	3.55E-05	2.31E-04	0.45

Key: FDR, False discovery rate

Chapter 4: Probe the mechanism of actions underlying cell fate change induced by metabolite

Based on manuscript draft in preparation for submission by **Jie Li**, Kaitlyn E. Stepler, Leena R. Pade, Camille Lombard-Banek, and Peter Nemes*

My contributions (the lead author): I designed the research; I developed and evaluated FACS-based sampling workflow; I collected samples, performed metabolomic analysis and processed the data; I conducted stable isotope labeling experiment and designed/performed functional studies; I analyzed the data, interpreted results, and wrote the manuscript draft.

4.1 Abstract

Metabolites are drivers, regulators, and products of a myriad of biological processes. Recently, metabolites were discovered to alter the normal developmental fate of embryonic stem cells. However, the mechanism of action underlying the metabolite-induced cell fate change remains unknown. This is largely due to the lack of techniques to isolate and analyze the progeny of the injected cells from developing embryos. To fill this knowledge gap, we developed a workflow that combines microinjection, cell lineage tracing, embryonic dissociation, and fluorescence-activated cell sorting (FACS) to track and isolate descendant cell clones, followed by metabolomic and proteomic analysis of isolated cell clones using liquid chromatography-high resolution mass spectrometry (LC-HRMS). A total of 55 metabolites and 1,851 proteins were identified with high confidence, which could be ascribed to metabolic pathways including the TCA cycle, glycolysis, arginine-proline metabolism, and cysteine and methionine metabolism. Statistical analyses of the quantitative ‘omic data revealed significant ($p<0.05$) metabolic and

proteomic changes as the cells' tissue fate changed. Applying the multi-omic strategy, I uncovered a molecular interaction network perturbed by injected methionine including upregulated the TCA cycle, glycolysis, and PPP and downregulated arginine-proline metabolism. The molecular interaction network suggested potential molecular effectors involved in cell fate change actions that were further validated by heavy isotope labeling experiments and functional studies. The revealed metabolic and proteomic changes suggested perturbation of a large metabolic network underlying the mechanism of action for the cell fate changes and provided new insights into cell fate decisions during early vertebrate embryonic development.

4.2 Introduction

While there is a wealth of knowledge about the role of genes, transcripts, and proteins in cell differentiation and tissue formation,[6, 167] we have little knowledge about the role of metabolites[6] (molecular weight < 1,500 Daltons) in these processes. Recent studies found metabolites to be actively involved in many functions outside of chemical reactions. Therefore, elucidating the role of metabolites can deepen our understanding of various biological processes including cell differentiation and tissue formation, and help to explore molecular mechanisms involved in these processes.[1, 37] The necessity to study the role of metabolites is further enhanced considering that metabolites have recently been identified as active drivers and modulators of functions of DNA, RNA, and proteins.[1, 168]

Most recently, the distribution of metabolites was found to be heterogenous between differentially fated cells in the 16-cell *X. laevis* frog embryos, and altering the known concentration of these metabolites caused cell fate change.[6] For example, methionine and acetylcholine were enriched in the epidermal-fated cells (called V11 cells), whereas serine and threonine were enriched in the neural-fated cells (called D11 cells).[6] Furthermore, these differentially enriched metabolites were demonstrated to cause predetermined cell fate change: the progeny of the D11 cells became epidermally fated when injected with methionine and acetylcholine.[6] However, the mechanism underlying the metabolite-induced cell fate changes remains unknown due to the lack of techniques to isolate and analyze tissue-specific cell clones descending from the initially injected blastomere in developing *X. laevis* embryos.

The development of sampling techniques amenable to ‘omic analysis has opened opportunities to study complex biological systems. Single cells can be manually dissected from the central nervous system to study the metabolic differences among different types of neurons in *Aplysia californica*.[42] Furthermore, single mammalian cells were penetrated by capillary microprobes to directly aspirate cellular contents for measuring small molecular peaks.[169] Moreover, microprobe sampling was further developed to enable *in vivo* multi-omic study of developing embryonic cells without perturbing normal embryonic development[26] and to study the subcellular distribution of peptides in single neurons[28]. However, the dissection- and aspiration-based sampling techniques have a relatively low throughput and require the dexterity of the researchers, thus, making them unamenable when sampling thousands of tissue-specific cells that can be dispersed during embryonic development.

Cell sorting techniques provide alternative ways to collect tissue-specific cells with labeled markers. Magnetic-activated cell sorting (MACS) is one sorting technique, in which cells having magnetic characteristics can be manipulated and sorted out via magnetic forces.[170] Magnetic particles are often conjugated with antibodies to selectively bind to a surface marker on the cells of interest, followed by sorting using a magnet to isolate these targeted cells. However, tissue-specific cells at early embryonic stages usually have no known surface markers, which precludes the development of corresponding antibodies, thus limiting the application of MACS for isolating tissue-specific cells. Fluorescence-activated cell sorting (FACS) is another sorting technique based on cell size, internal structure, and fluorescence profile.[171] FACS is widely used as it allows the sorting of specific types of cells with high purity and viability. For example, FACS was applied to isolate muscle stem cells from adult mice,[172] vascular cells from human embryonic stem cells,[173] and circulating tumor cells.[174] FACS combined with a recently developed dissociation strategy[79] inspires new sampling strategies for isolating tissue-specific cells that would be difficult to dissect or microsample from developing organisms.

Characterizing different types of biomolecules provides complementary pathway enrichment results and expands metabolic pathway coverage[26] but faces analytical challenges. Metabolomic analysis of tissue-specific cell clones provides a snapshot of small molecules while proteomic analysis supplies information about proteins, including enzymes involved in the production of various metabolites, to together delineate the state of cells. In addition to protein identification and quantification, proteomics can also reveal post-translational modifications (PTMs) on proteins that can significantly impact their functions.[175] However, such analyses are challenging in tissue-specific cell clones due to the limited amounts of starting materials and the fact that both metabolites and proteins are not amplifiable. Furthermore, metabolites have a wide physiological concentration range,[48] can change rapidly,[48] and vary in physicochemical properties[37]. In addition, there are more than 114,000 metabolites to be considered based on Human Metabolome Database[137] and these metabolites are structurally diverse[48]. Protein concentrations also cover a wide dynamic range,[176-177] which can be further complicated by the presence of high-abundance proteins that mask the detection of biologically relevant low-abundance proteins.[22, 178] All of these factors together increase the complexity of detecting, identifying, and quantifying metabolites and proteins from biological systems.

MS is widely used for metabolomic and proteomic analysis because it allows for detecting large numbers of biomolecules simultaneously from limited samples.[4, 12, 37-38, 48, 179] MS is a label-free technique that offers high sensitivity, wide dynamic response range, and structural specificity. Coupling separation techniques such as capillary electrophoresis (CE) and liquid chromatography (LC) to MS not only decreases the complexity of ‘omic analysis to facilitate biomolecule identification[180] but also adds separation time as another dimension for confident metabolite identification.[36] Furthermore, LC-MS has the advantage of being able to tailor to the various physicochemical properties of biomolecules to drive separation. For example, the combination of high- and low-pH reversed-phase LC leads to a complementary separation of peptides and a corresponding improvement in protein identifications.[44] Recently, advances in LC separation for proteomics such as the development of the micro-pillar array column (μPAC) have further improved peptide separation, reproducibility, and stability compared to traditional packed-based LC columns to facilitate deeper proteome coverage from limited biological samples.[181]

Here, to uncover the mechanisms of action underlying the metabolite-induced cell fate change, we developed a sampling workflow to isolate tissue-specific cells by integrating microinjection, cell lineage tracing, embryonic dissociation, and FACS (**Fig. 1**). The collected cell clones were processed for dual metabolomic and proteomic analysis using LC-HRMS which identified 55 metabolites and 1,851 proteins with high confidence, of which 16 metabolites and 279 proteins displayed statistically significant change (*p*-value <0.05) by injected methionine. When mapping these detected metabolites and proteins to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database[165], downregulated arginine-proline metabolism and upregulated TCA cycle, glycolysis, and pentose phosphate pathway were uncovered. The molecular interaction network involving these pathways showed potential molecular effectors responsible for the cell fate change actions, which were further validated by stable isotope labeling experiments and functional studies. Our results revealed molecular remodeling underlying metabolite-induced cell fate changes, which suggested a hypothesis that the injected methionine upregulated the level of carnitine and acetylcarnitine, which further perturbed the level of acetyl-CoA. The perturbed acetyl-CoA level affected energy production and acetylation PTMs of proteins, resulting in cell fate change. This study shed light on a long-standing biological question to explore the role of metabolites in driving cell differentiation.

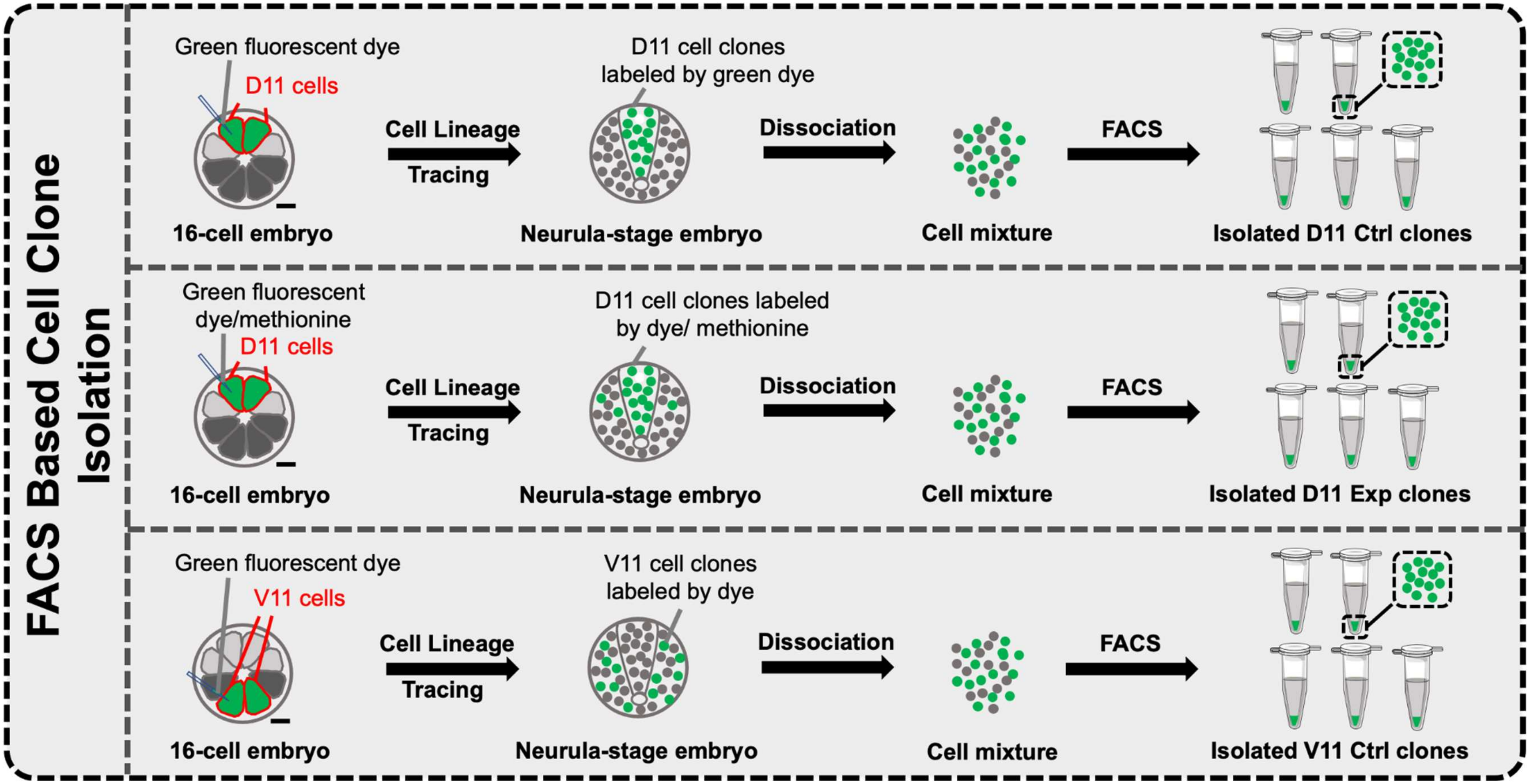


Figure 4.1 FACS-based tissue-specific cell clone isolation by combining microinjection, cell lineage tracing, embryonic dissociation, and FACS. 3 types of cell clones at neurula stage 13 were collected using this sampling strategy. D11 cell clones labeled with green fluorescent dye only, called D11 Ctrl (top); D11 cell clones labeled with green fluorescent dye and methionine, called D11 Exp (middle); and V11 cell clones labeled with green fluorescent dye, called V11 Ctrl (bottom). Scale bar: 250 µm.

4.3 Experimental section

4.3.1 Materials and chemicals

Dextran Alexa Fluor 488 dye from Thermo Fisher Scientific (Invitrogen D34682). Bovine serum albumin (BSA), acetylcholine, methionine, threonine, arginine, and guanine were from Fisher Scientific (Pittsburgh, PA). Sodium isethionate, sodium pyrophosphate, sodium butyrate, CAPS, proline, L-methionine-$^{13}C_5$,^{15}N, and adenosine were from Sigma Aldrich (St. Louis, MO). Citrulline, cysteine, and isoleucine were from MP Biomedicals (Solon, OH). All solvents were LC-MS grade unless specified and were from Fisher Scientific (Pittsburgh, PA). All chemicals were reagent grade or higher unless specified.

4.3.2 Solutions

2% (w/v) cysteine solution (pH 8). 100% Steinberg's solution (SS, containing 58.2 mM NaCl, 0.67 mM KCl, 0.34 mM $Ca(NO_3)_2$, 0.83 mM $MgSO_4$, 4.19 mM Trizma hydrochloride, and 0.66 mM Trizma base, in Milli-Q purified water). 3% (w/v) Ficoll solution in 100% SS. 50% SS. 1 mg/mL BSA solution. Newport 10.5 dissociation buffer (containing 100 mM sodium isethionate, 20 mM sodium pyrophosphate, and 10 mM CAPS, pH 10.5, in Milli-Q purified water). 1× Marc's Modified Ringers (MMR) buffer (containing 100 mM NaCl, 2 mM KCl, 1 mM $MgSO_4$, 2 mM $CaCl_2$, 5 mM HEPES, and 0.1 mM EDTA, pH 7.8, in Milli-Q purified water). 1× phosphate-buffered saline (PBS) buffer. 1× PBS buffer with 2% BSA. 4% paraformaldehyde in 1× PBS. 0.5% (w/v) Dextran Alexa Fluor 488. 0.5% Dextran Alexa Fluor 488 with 50 mM methionine. 100 pg/nL gfp mRNA. 100 pg/nL gfp mRNA with 50 mM acetylcarnitine. 100 pg/nL gfp mRNA with 50 mM serine and threonine. 100 pg/nL gfp mRNA with 5 mM acetyl-CoA. 100 pg/nL gfp mRNA with 5 mM sodium butyrate. Lysis buffer (containing 1% sodium dodecyl sulfate (SDS), 150 mM NaCl, 5 mM ethylenediaminetetraacetic acid (EDTA), 20 mM Tris-HCl). Digestion buffer (containing 50 mM ammonium bicarbonate adjusted to pH 8 with Tris).

4.3.3 X. laevis frog care and embryo preparation

X. laevis animal care and handling followed the requirements and regulations of the Institutional Animal Care and Use Committee (IACUC) of the University of Maryland (IACUC No. R-FEB-21-07). Embryos were obtained via natural mating induced by human chorionic gonadotropin (hCG).[141] The obtained embryos were removed of jelly coating using 2% cysteine (pH 8).[141] 2-cell embryos displaying stereotypical cleavage and pigmentation were cultured in 100% SS under 14 °C until the 16-cell stage for microinjection.[141]

4.3.4 Needle preparation, calibration, and embryo injection

Borosilicate glass capillaries (Sutter Instrument, O.D./I.D.: 1.0 mm/0.50 mm) were pulled by a micropipette puller (P-1000, Sutter Instrument) using the following settings: heat = 495; pull = 30; velocity = 60; time = 150; pressure = 200. Pulled needles were cut off 60 units using sharp forceps under 40× magnification, filled with 1 µL injection solution, and calibrated to ensure 1 nL per injection following the established procedure[141]. The injection solution was 0.5% Dextran Alexa Fluor 488 for the control group while the injection solution was 0.5% Alexa Fluor 488 with 50 mM methionine for the experimental group. 16-cell stage embryos showing stereotypical cleavage and pigmentation were placed in an injection plate filled with 3% Ficoll in 100% SS for microinjection. For the negative control group, both the left and right D11 cells were injected with 1 nL of 0.5% Dextran Alexa Fluor 488. Similarly, both the left and right V11 cells were injected with 1 nL of 0.5% Dextran Alexa Fluor 488 for the positive control group. In contrast, both the left and right D11 cells were injected with 1 nL of 0.5% Dextran Alexa Fluor 488 with 50 mM methionine for the experimental group.

4.3.5 Cell lineage tracing and embryonic dissociation

Both injected and uninjected embryos were raised to stage 13 in 50% SS. D11 and V11 cell clones were tracked and recognized using a fluorescence stereomicroscope. Embryos at stage 13 were removed of the vitelline membrane first, followed by washing with ice-cold 1× MMR solution, and transferred to a BSA-pretreated 12-well plate (Corning 3513) filled with ice-cold Newport 10.5 dissociation buffer.[79] For the best dissociation efficiency, each well contained no more than 6 embryos. To minimize cell lysis during dissociation, wells were sealed with parafilm to remove air. The sealed plates were vortexed (Genie2, Scientific Industries) for 5 min on setting 3, followed by 15 min vortexing on setting 5 and 5 min vortexing on setting 7. Embryos without injection were dissociated simultaneously.

4.3.6 Cell suspension preparation and FACS

The 12-well plates after dissociation were placed on ice for 10 min to let cells settle down. The dissociation solution was removed as much as possible using a transfer pipette, followed by adding 400 μL of 2% BSA in 1× PBS to each well immediately. The obtained cell suspensions were transferred to 1.5 mL Eppendorf vials using a 1 mL Eppendorf pipette. To minimize sample loss, each well was rinsed using 400 μL of 2% BSA in 1× PBS and the rinse solutions were combined with the obtained cell suspensions in the Eppendorf vials. Afterward, vials were centrifuged at 4°C, 400 ×g for 5 min to form cell pellets. The supernatants were discarded to remove the dissociation solution, and the pellets were added with 400 μL of 2% BSA in 1× PBS. Sample vials were stored in an ice bath until FACS experiments. Cells were resuspended by gently pipetting up and down using a 1 mL pipettor right before the FACS experiment. The obtained cell suspensions were filtered using a 40 μm cell strainer (542040, Greiner Bio-One), followed by transferring to 12×75 mm tubes for sorting. Cell suspensions from uninjected embryos were processed in the same way and used as controls to determine the level of background fluorescence and autofluorescence. Cell suspensions from injected embryos were sorted based on cell size and fluorescence profiles. Both the fluorescently labeled and unlabeled cells were collected in 12×75 mm tubes pre-filled with 600 μL of 2% BSA in 1× PBS in ice-cold baths. The collected cell suspensions were transferred to 1.5 mL Eppendorf vials, followed by centrifuging for 5 min at 4°C and 400 ×g. The supernatant was discarded, and the obtained cell pellets were added with 100 μL ice-cold LC-MS grade methanol to quench metabolism. Samples were stored in a −80°C freezer immediately until metabolite extraction. Using the developed sampling method, we collected 3 groups of samples including D11 Ctrl (from precursor D11 cells injected with green fluorescence dye only), D11 Exp (from precursor D11 cells injected with green fluorescence dye and methionine), and V11 Ctrl (from precursor V11 cells injected with green fluorescence dye only). Each group was collected with 5 or 6 biological replicates to ensure sufficient confidence in statistical analysis.

4.3.7 Metabolite extraction and HILIC-ESI-MS analysis

The collected cells were taken from the −80°C freezer and dried at 4°C using a vacuum concentrator (Labconco). 400 μL of a mixture of 40% (v/v) acetonitrile and 40% (v/v) methanol in water were added to extract metabolites. Samples were vortexed for 30 s, and sonicated for 10 min in an ice bath, followed by 10 min incubation in the ice bath. This vortexing-sonication-incubation process was repeated twice. Afterward, the samples were incubated in a −20°C freezer for 1 h to precipitate proteins, followed by centrifuging for 10 min at 4°C and 13,000 ×g. The supernatants were transferred to an Eppendorf vial and vacuum dried at 20°C. The residues were reconstituted in 50 µL 95% (v/v) acetonitrile in water and centrifuged for 10 min at 4°C and 13,000 ×g. The supernatants were transferred to HPLC vials for HILIC-ESI-MS analysis using a Waters ACQUITY UPLC I-Class system coupled to a time-of-flight (TOF) mass spectrometry (MS) (Bruker timsTOF Pro). The samples were queued in an autosampler (4°C) in randomized order and injected with 2 µL. Metabolites were separated by a Waters HILIC column (ACQUITY BEH Amide, 1.7 µm, 1 mm×100 mm) (40°C) at a flow rate of 130 µL/min under the following LC gradient: 0–0.5 min 95% B, 0.5–10min 95–40% B, 10–13 min 40% B, 13–15 min 40–95% B, 15–22 min 95% B. Mobile phase is: A = 0.1% formic acid (FA) in water (LC-MS grade) and B = 0.1% FA in acetonitrile (ACN, LC-MS grade). LC-MS data were acquired under the positive mode using an ion booster source (End Plate Offset: 500 V; capillary 4,500 V; nebulizer: 0.4 bar; dry gas: 4 L/min; dry temperature: 220°C): mass range, *m/z* 20–1,000; spectra rate, 2 Hz. MS/MS experiments were performed using automated data-dependent acquisition (DDA) with a spectra rate of 10 Hz and a threshold of 1,000 counts. Each biological replicate was measured with 3 technical replicates.

4.3.8 Proteomic sample preparation and nanoLC-ESI-MS analysis

After metabolite extraction, the pellet for each sample was resuspended in 10 µL lysis buffer and aspirated several times to lyse cells and extract proteins. Extracted proteins were reduced with 0.4 µL 0.5 M dithiothreitol (DTT) at 60°C for 30 min and alkylated with 1.2 µL 0.5 M iodoacetamide for 20 min in the dark. Alkylation was quenched with the same volume of DTT, and samples were vortexed immediately. Protein was precipitated with chilled acetone overnight at –20°C and pelleted via centrifugation at 10,000 ×g at 4°C for 10 min. The supernatants were removed to waste and protein pellets were reconstituted in a digestion buffer. Trypsin (0.75 µg) was added to each sample and proteins were digested at 37°C for 12 h. The resulting peptides were dried by centrifugal evaporation and quantified using a Take3 micro-volume plate (Agilent Technologies, Santa Clara, CA). A pooled sample was created from equimolar peptide amounts of all samples for quality control (QC) purposes. One pooled sample was labeled as part of each batch of samples with Tandem Mass Tags (TMT)^{10}-plex (Thermo Fisher Scientific, Waltham, MA; 25 μg peptides per sample). Samples were randomly assigned to TMT reagent channels. Each sample was reconstituted to 1 μg/μL in 100 mM triethylammonium bicarbonate and incubated with 0.2 mg of TMT for 1 h with vortexing for labeling. The reaction was quenched with 2 μL 5% hydroxylamine and incubated for 15 min with vortexing. After labeling, the samples from each batch were combined and dried by centrifugal evaporation. Each batch mixture (100 μg) underwent offline high pH reversed-phase fractionation into 72 fractions using an Agilent 1260 Infinity II liquid chromatography (LC) system coupled with a fraction collector (Agilent). TMT-labeled peptide mixtures were loaded onto an InfinitiLab Poroshell HPH-C18 guard column (Agilent; 4.6×5 mm, 2.7 μm) and then separated on an InfinitiLab Poroshell HPH-C18 column (Agilent; 4.6×150 mm, 2.7 μm) at 0.5 mL/min over a gradient of 0–17 min, 0% B; 17–18 min, 0–7% B; 18–75 min, 7–35% B; 75–76 min, 35–100% B; 76–80 min, 100% B; 80–86 min, 100–0% B; 86–90 min, 0% B. Mobile phases A and B were 10 mM ammonium bicarbonate in water and 90% ACN, respectively, both adjusted to pH 10 with ammonium hydroxide. Fractions were collected between 17–85 min, dried via centrifugal evaporation, and concatenated to 8 final fractions. Each fraction was reconstituted in 0.1% FA for LC-MS^3 analysis. Fractions were injected in technical triplicate (1 μg each) for LC-MS^3 analysis on an UltiMate™ 3000 RSLCnano system (Thermo) coupled to an Orbitrap Fusion Lumos mass spectrometer (Thermo) operated in positive mode. Fractions were analyzed in random order. Peptides were loaded onto an Acclaim PepMap™ 100 C18 trap column (Thermo, 100 µm×2 cm, 100 Å, 5 µm) at 5 µL/min and then were separated on a µPAC™ column (Thermo; 200 cm) at 0.600 µL/min and 50°C over a 240 min gradient as follows: 0–5 min, 1% B; 5–20 min, 1–7% B; 20–135 min, 7–25% B; 135–160 min, 25–32% B; 160–193 min, 32–45% B; 193–200 min, 45–75% B; 200–208 min, 75% B; 208–210 min, 75–2% B; 210–240 min, 2% B. Mobile phases A and B were 0.1% FA in LC-MS grade water and ACN, respectively. Full MS spectra were collected in the Orbitrap (400–1,600 *m/z*, 120,000 resolution). MS/MS scans were collected in the ion trap in DDA mode with a 2.5 s cycle time using collision-induced dissociation (CID; 35% normalized collision energy, 0.7 *m/z* isolation window, turbo ion trap scan rate) with a dynamic exclusion of 60 s. The peptide monoisotopic peak determination node was on and only precursor ions were required to have charges +2–+6 and an intensity >5.0×10^3 to be considered for MS/MS fragmentation. MS^3 spectra were collected in the Orbitrap using synchronous precursor selection (SPS) to simultaneously fragment the top 10 most intense fragments from the MS/MS spectra using higher-energy C-trap dissociation (HCD; 100–500 *m/z*, 55% normalized collision energy, 50,000 resolution, 0.7 *m/z* MS isolation window, 3 *m/z* MS/MS isolation window). Precursor ions for MS^3 were selected from the 400–1,600 *m/z* range with a ±25 ppm exclusion mass width and precursor ions with a TMT tag loss were excluded. All spectra were collected using the standard automatic gain control (AGC) target and auto maximum injection time mode.

4.3.9 Data processing

The obtained metabolomics MS data files were processed using Bruker Compass MetaboScape Version 4.0.4 to find molecular features (unique *m/z* and retention time) along with peak intensity for quantitative analysis using the following setting: intensity threshold (1,000 counts); minimum peak length (10 spectra). Metabolite identification was based on accurate mass, isotopic pattern, and MS/MS matching[182] against the with METLIN database[11] and MassBank of North America (MoNA)[183] databases. The identification tolerance and scoring were 10 ppm for *m/z*, 10–50 for mSigma, and 400–900 MS/MS matching score. Metabolite standards were used for further confirmation of selected identified metabolites. The identified metabolites were listed in Supplementary Table S1. MetaboAnalyst 5.0[160] was used to perform statistical analyses to identify metabolites with statistically significant changes (p-value < 0.05) caused by injected methionine into D11 cell clones. Pathway analyses were conducted using MetaboAnalyst 5.0 to map detected metabolites to known metabolic pathways on the KEGG database.

RAW files for proteomics data were analyzed using Proteome Discoverer software (version 2.2) with all fractions and technical replicates for each batch combined into one file for analysis. Each batch file was searched against a concatenated database of *X. laevis* proteins including proteins from Xenbase[139] and the mRNA-derived Proteomic Reference with Heterogeneous RNA Omitting the Genome (PHROG[162]), as well as a database of common contaminants, using SEQUEST-HT. Dynamic modifications included in the search were TMT^{10}-plex (229.163 Da) on peptide N-termini and lysine residues, methylation of lysine, histidine, and arginine residues, oxidation of methionine residues, and acetylation of the protein N-termini as well as the fixed modification of carbamidomethylation on cysteine residues. The search allowed two trypsin missed cleavages and used mass tolerances of 10 ppm for precursor ions and 0.6 Da for fragment ions. Only peptides with a false discovery rate (FDR) of ≤1% were included.

The TMT^{10}-plex quantification method was used with corrections for the isotopic impurities of the TMT lots used in these experiments. Reporter ions were quantified by intensity with a signal-to-noise ratio of ≥30. Reporter ion intensities were normalized to the total peptide amount. Protein identifications in each batch were filtered to only include Master Proteins with FDR ≤1% and ≥2 peptide spectral matches (PSMs), and to exclude contaminant proteins. Quantified proteins were identified in all samples and had TMT reporter ion intensities for all pooled samples and ≥3 biological replicates per group. The intensities for quantified proteins were $\log_{10}$ transformed and used to calculate fold changes and *p*-values between sample groups using two-tailed Student's *t*-tests. Fold-change cutoffs were calculated for each group comparison based on technical and biological variation and replication[184]. Differentially expressed proteins had $p < 0.05$ and met the defined fold-change cutoffs. Interactions among differentially expressed proteins were predicted using STRING version 11.5[185]. Gene Ontology annotation and statistical overrepresentation analyses were performed for differentially expressed proteins against *Xenopus tropicalis* using the PANTHER database[186] version 17.0 (released 02/22/2022).

4.3.10 Heavy isotope labeling to track the metabolic fate of methionine

16-cell embryos displaying stereotypical cleavage and pigmentation cultured from 2-cell embryos were selected for microinjection. To track the fate of methionine, both left and right D11 cells on 16-cell embryos were microinjected with 0.5% Alexa Fluor 488 containing 250 pmol heavy-isotope labeled methionine-$^{13}C_5$,^{15}N. The injected embryos were cultured to neurula stage 13, followed by the dissection of fluorescently labeled D11 cell clones under a stereomicroscope (SMZ18, Nikon) from 10 embryos. The dissected D11 cell clones underwent metabolite extraction to use the supernatant layer for metabolomic analysis as mentioned above. The extracted residue (pellet) was processed for bottom-up proteomic analysis as mentioned above, except using a label-free approach instead of TMT quantification and without offline fractionation. After digestion, resulting peptides were desalted using C18 spin columns

(Thermo), dried via centrifugal evaporation, and reconstituted in 0.1% FA for LC-MS/MS analysis. LC-MS/MS analysis was the same as the above LC-MS^3 analysis with the following differences: full MS scan range 350–1,500 *m/z* and maximum injection time 50 ms, cycle time 3 s, +2–+7 charge states allowed, HCD-MS/MS spectra collected in the ion trap with 30% normalized collision energy, 1.6 *m/z* isolation window, rapid ion trap scan rate, auto scan range mode, and dynamic maximum injection time mode.

4.3.11 Functional studies and image scoring

To validate whether the revealed upregulated metabolites resulted in cell fate change of neural precursor cells, right D11 cells in 16-cell embryos were microinjected with 1 nL of 100 pg/nL *gfp* mRNA only for the control group. For the experimental groups, right D11 cells were microinjected 1 nL of 100 pg/nL *gfp* mRNA along with 50 mM acetylcarnitine, 50 mM serine and threonine, 5 mM acetyl-CoA, or 5 mM sodium butyrate, respectively. The injected embryos were cultured under room temperature (21°C) in 50% SS to larvae stage 34, followed by fixation by 4% paraformaldehyde in 1× PBS for 1 hour and washing with 1× PBS for 15 minutes. The fixed larvae were imaged using a stereomicroscope (SMZ18, Nikon) for brightfield and green fluorescence images. The obtained images were scored for neural and epidermal tissues following a previous protocol: score 0 for tissues showing no fluorescently labeled cells; score 1 for tissues showing ≤10 labeled cells; score 5 for tissues showing many labeled cells; score 10 for tissues almost completely labeled.[6] Mann-Whitney test was performed to evaluate whether there were statistically significant changes of tissue scoring induced by injected metabolites.

4.4 Results and discussion

To elucidate the molecular mechanisms underlying metabolite-induced cell fate, I developed a workflow that combines microinjection, cell lineage tracing, embryonic dissociation, and FACS to track and isolate descendant cell clones of injected precursor cells followed by LC-MS-based metabolomic and proteomic analysis (**Fig. 4.2**). Pathway and statistical analyses were performed to reveal perturbed metabolic/proteomic changes by injected metabolites, which provided insight into the molecular mechanisms of cell fate change actions.

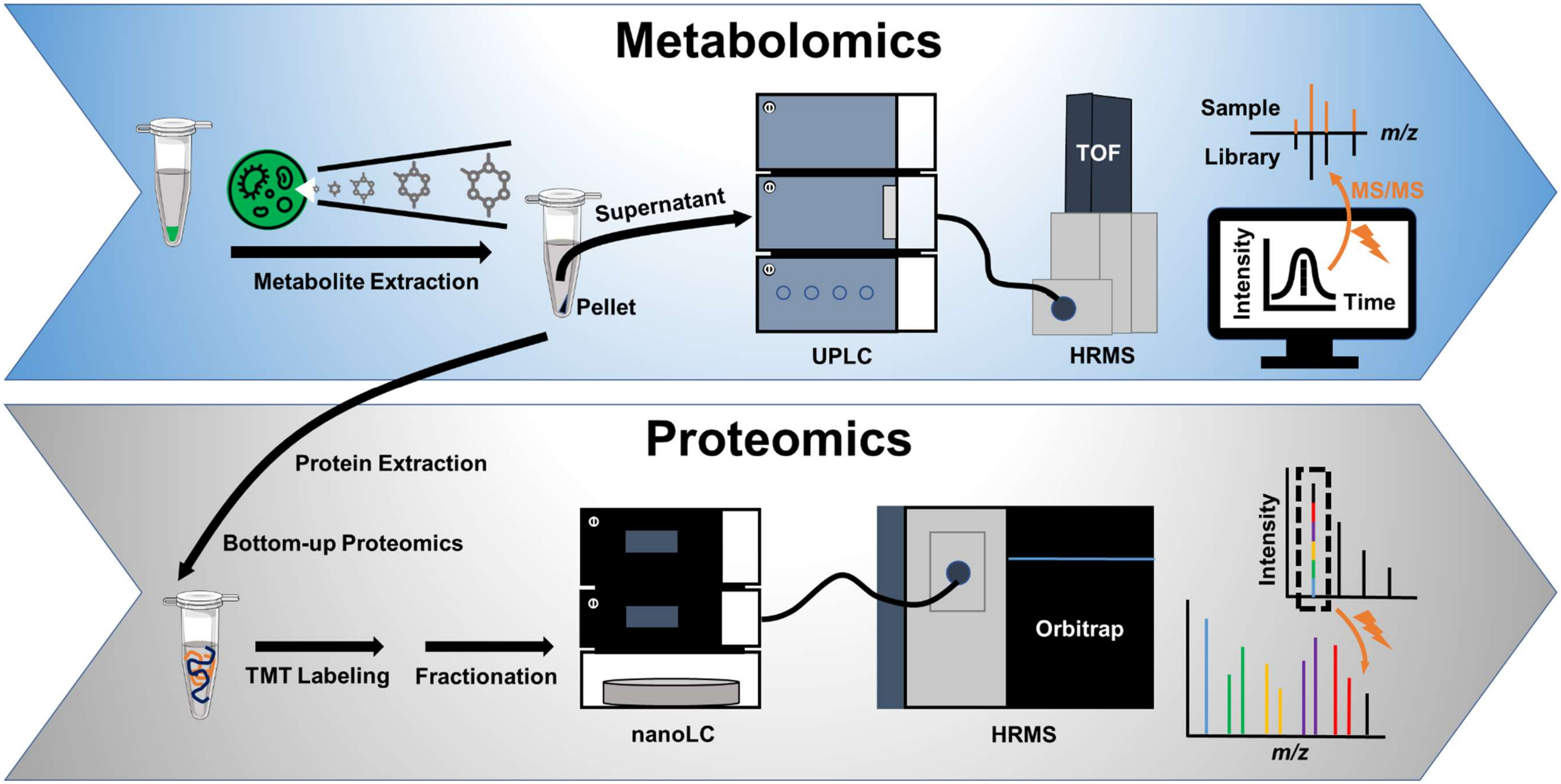


Figure 4.2 Simultaneous isolation of metabolites and proteins for LC-HRMS-based metabolomics and proteomics of isolated cell clones.

4.4.1 Developed and evaluated FACS-based sampling workflow for tissue-specific cell clones isolation

I demonstrated microinjection combined with cell lineage tracing, embryonic dissociation, and FACS was feasible to label and isolate tissue-specific cell clones from neurula-stage 13 *X. laevis* frog embryos. To label and track tissue-specific cell clones, I microinjected (**Fig. 4.3A**) green fluorescent dye into precursor cells in 16-cell stage embryos (**Fig. 4.3B**), followed by fluorescence microscopy to track the fluorescently labeled cell clones over development at later stages. **Fig. 4.3B** showed D11 cell clones labeled by fluorescent dye. By conducting embryonic dissociation, we successfully obtained a cell mixture of fluorescently labeled and -unlabeled cells from the injected embryos (**Fig. 4.3C**). To test the cell viability of dissociated cells, I performed the trypan blue exclusion tests[187], which showed that 95.6% of dissociated cells were viable. Further investigation of dissociated cells under a microscope, I found cells from neurula stage 13 embryos had diameters ranging from 10 to 35 μm, which were compatible with workable cell sizes by general FACS instruments. By using FACS, the green fluorescently labeled cells—targeted cell clones were collected based on cell size, internal granularity, and fluorescence profile (**Fig. 4.3D**). **Fig. 4.3E** was an example image of the FACS collected cell clones which showed all the cells in the image were fluorescently labeled. Statistical analysis of multiple fluorescent images of collected cells showed that the purity of FACS-sorted cell clones was 95.1% (192/202). Furthermore, the collected cells were determined to have cell viability of 92.9% according to the trypan blue tests. To check the potential sample loss during FACS, I

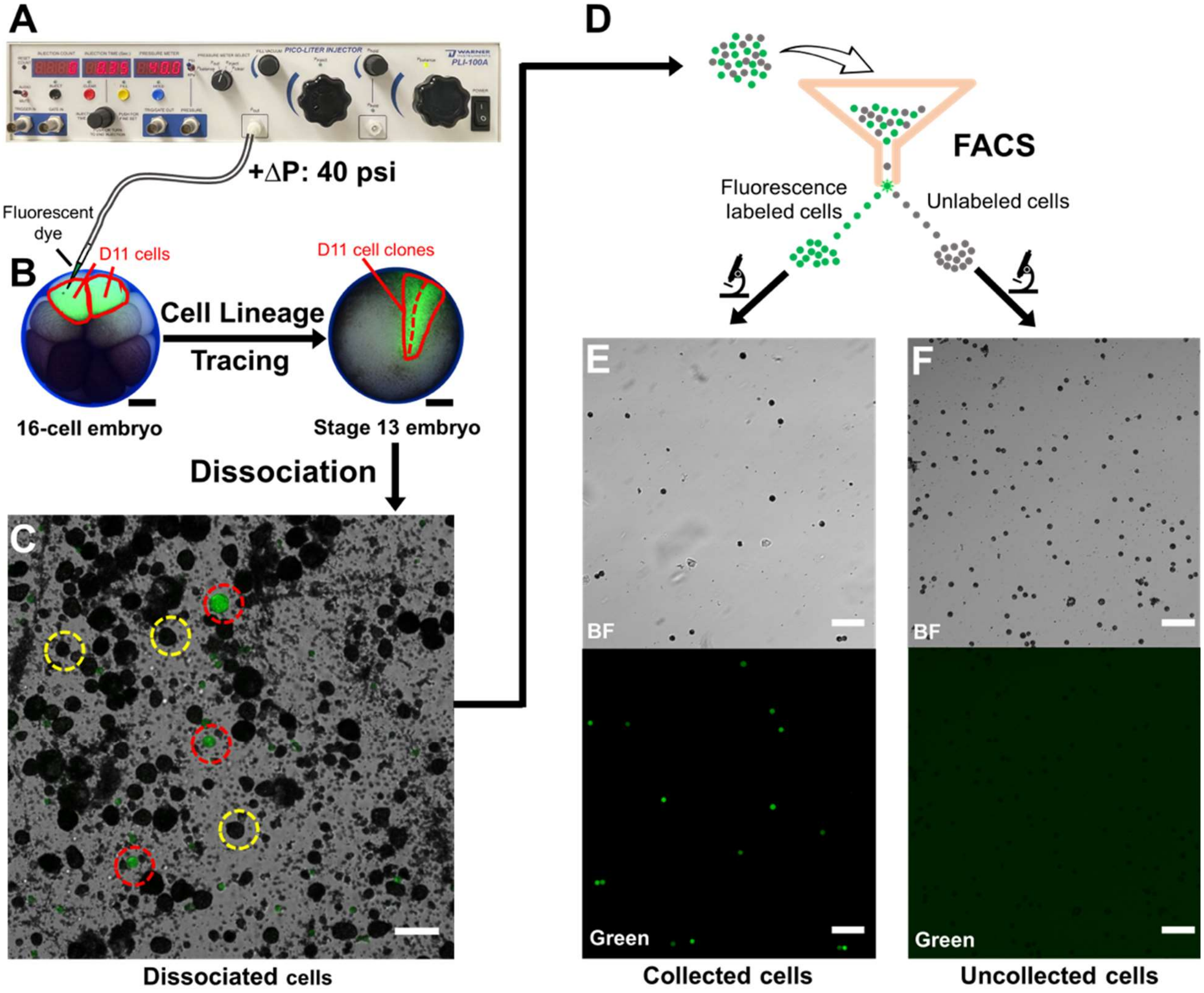


Figure 4.3 FACS-based workflow to isolate targeted cell clones using microinjection, cell lineage tracing, embryonic dissociation, and FACS. (**A**) Microinjector used to inject green fluorescent and/or methionine into both left and right D11 cells. (**B**) Cell lineage tracing under a fluorescence microscope to track the cell clones of injected cells. Left and right D11 cells in 16-cell embryos were labeled by green fluorescent dye as highlighted in red; neurula stage 13 embryo with D11 cell clones labeled by green fluorescent dye was highlighted in red. Scale bar: 250 µm. (**C**) Cell mixture from dissociated neurula stage 13 embryos. Cell circled in red color are D11 cell clones. Cell circled in yellow color are non-D11 cell clones. Scale bar: 50 µm. (**D**) Illustration of FACS to isolated fluorescently labeled cell clones. (**E**) Images of D11 cell clones sorted during FACS. Top: brightfield (BF) image; Bottom: green fluorescent image. Scale bar: 200 µm. (**F**) Images of cells went to waste during FACS. Top: brightfield (BF) image; Bottom: green fluorescent image. Scale bar: 200 µm.

collected and imaged cells that went to waste during FACS. The fluorescence image showed that there was barely any sample loss (**Fig. 4.3F**). In summary, I established a workflow to isolate tissue-specific cell clones from *Xenopus* neurula-stage embryos with high purity, high viability, and minimal sample loss.

To test if the molecular changes happen at the early embryonic stage perturbed by injected methionine, we focused on biologically important Nieuwkoop and Faber (NF) stage 13. This is because NF stage 13 is the stage right after gastrulation in which cells migrate to form endoderm, mesoderm, and ectoderm—developing into different tissues and organs later. Moreover, NF stage 13 is the starting stage of the neurula stage in which the neural tube is formed and further develops into the central nervous system over embryonic development. In terms of technical capability, the size of cells in the neurula stage is small enough to be compatible with the general FACS instrument. Applying the developed FACS-based sampling workflow to study metabolite-induced cell fate change at the neurula stage, I achieved the isolation of both D11 cell clones and V11 cell clones (**Fig. 4.1**).

4.4.2 Established a HILIC-HRMS method for metabolomic analysis

To detect and quantify metabolic changes happening as cell fate changed, I developed a sensitive and robust HILIC-HRMS method by using mixed metabolite standards. According to the calibration curve of metabolite standards (**Fig. 4.4**), the developed HILIC-HRMS method has a wide dynamic linear response range (~4 orders of magnitude). By calculating the signal-to-noise ratio (S/N), I found the HILIC-HRMS method was sensitive enough to detect low-abundant metabolites (0.5 nM adenosine, S/N=17.9; 5 nM isoleucine, S/N=5.8; 10 nM methionine, S/N=20.3; 10 nM threonine, S/N=22.5). I also evaluated the technical reproducibility

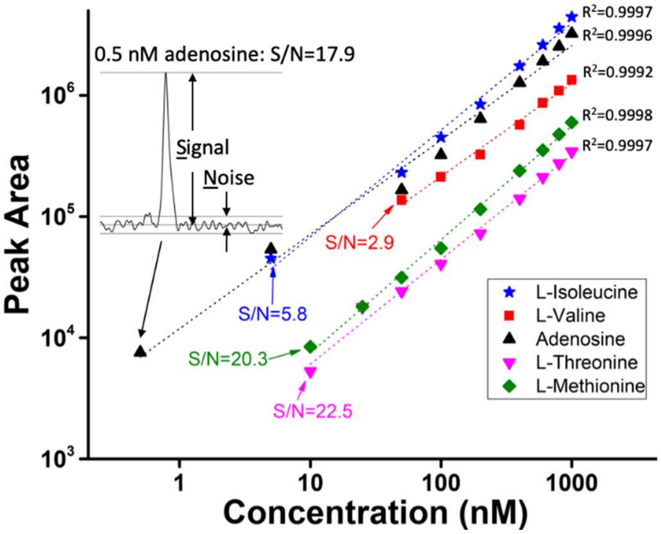


Figure 4.4 Calibration curve of 5 metabolite standards using the developed HILIC-HRMS. The inserted figure shows the calculation of S/N of 0.5 nM adenosine.

of the developed HILIC-HRMS method, which achieved a reproducibility of <0.5% relative standard deviation (RSD) for retention times and <16% RSD for peak areas. In summary, the developed HILIC-HRMS method was sufficient for the detection and quantification of low-abundant metabolites from isolated cell clones. I performed untargeted metabolomics for metabolite extract from collected D11 Ctrl, D11 Exp, and V11 Ctrl groups using the developed ultra-sensitive HILIC-HRMS method, which enabled the detection of 3,763 molecular features and the identification of 55 metabolites with high confidence based on accurate mass, retention time matching, and MS/MS matching with the METLIN database[11].

4.4.3 Statistical analyses to reveal metabolic perturbations induced by injected methionine

Statistical analyses were performed to reveal metabolites that changed significantly (*p*-value <0.05) among D11 Ctrl, V11 Ctrl, and D11 Exp. Based on the volcano plot between D11 Ctrl (N = 6) and V11 Ctrl (N = 5), 21 of 55 identified metabolites were enriched differently (*p*-value <0.05) (**Fig. 4.5**). Specifically, 10 metabolites including guanine, hypoxanthine, acetylcarnitine, and tyrosine, and creatinine were enriched in V11 Ctrl while 11 metabolites including diethanolamine, isoleucine, glycerophosphocholine, and ornithine were enriched in D11 Ctrl. Moreover, the volcano plot between D11 Ctrl (N = 6) and D11 Exp (N = 5) showed that 16 of 55 metabolites were enriched differently (*p*-value < 0.05) (**Fig. 4.5**). Specifically, 12 metabolites including threonine, arginine, isoleucine, proline, serine, and ornithine were downregulated by injected methionine in D11 Exp while 4 metabolites including acetylcholine, carnitine, acetylcarnitine, and ectoine were upregulated in D11 Exp.

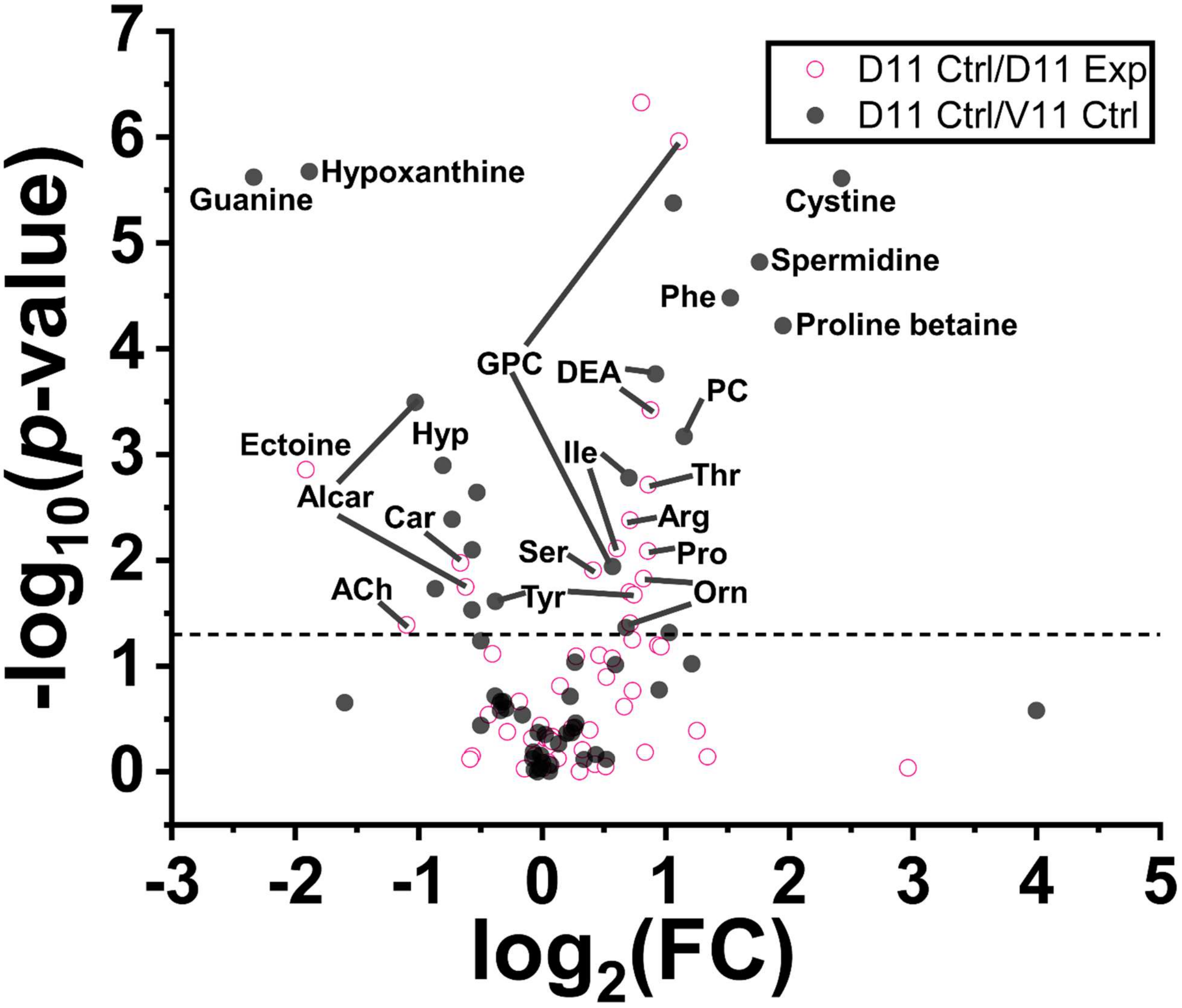


Figure 4.5 Statistical analysis of 3 different groups: neurula stage 13 D11 cell clones called D11 Ctrl, neurula stage 13 V11 cell clones called V11 Ctrl, and neurula stage 13 D11 cell clones with methionine called D11 Exp. Volcano plot shows metabolites displaying statistically significant differences (*p*-value <0.05) and the correspondingly fold change (FC). D11Ctrl/D11 Exp (empty red circle); D11Ctrl/V11 Ctrl (dark grey circle). Abbreviation: glycerophosphocholine (GPC), diethanolamine (DEA), threonine (Thr), arginine (Arg), isoleucine (Ile), proline (Pro), serine (Ser), ornithine (Orn), tyrosine (Tyr), carnitine (Car), acetylcarnitine (Alcar), acetylcholine (ACh), 4-hydroxyproline (Hyp), PC(18:1(9Z)/18:1(9Z)) (PC), phenylalanine (Phe).

When correlated with the volcano plots together, glycerophosphocholine, diethanolamine, isoleucine, and ornithine were found to be enriched in D11 Ctrl compared with both D11 Exp and V11 Ctrl (**Fig. 4.5**). Instead, both D11 Exp and V11 Ctrl were enriched with acetylcarnitine. This means that injected methionine caused D11 cell clones at the neurula stage to change their metabolic state toward V11 cell clones during embryonic development. To further explore the correlation between metabolite enrichment abundance and cell tissue fate over embryonic development, I compared the abundance differences observed between D11 Ctrl and D11 Exp at stage 13 with those between D11 cells and V11 cells at the 16-cell embryonic stage. I found that proline, threonine, ornithine, serine, and tyrosine had lower abundance in D11 Exp compared to D11 Ctrl at stage 13 (**Fig. 4.5**). The expression pattern of these 5 metabolites matched exactly with the abundance difference between D11 and V11 cells at 16-cell embryonic stage—V11 cells had a lower abundance compared to D11 cells[6]. Similarly, I found that D11 Exp had a higher abundance of acetylcarnitine and acetylcholine compared with D11 Ctrl at stage 13, which also paired with the same trend in V11 cells compared to D11 cells at the 16-cell stage[6]. This was exciting because it revealed that D11 Exp had methionine-induced metabolic changes at neurula stage 13 resembled the abundance difference at the 16-cell stage.

4.4.4 Statistical analyses to reveal proteomic perturbations induced by injected methionine

To complement the insight revealed by our metabolomic analyses, we also applied proteomics to investigate molecular changes induced by injected methionine. As for the metabolomics analyses, samples from D11 Ctrl (N = 5), V11 Ctrl (N = 4), and D11 Exp (N = 4) groups were analyzed by nanoLC-MS^3 using TMT for protein quantification. These analyses identified a total of 1,851 proteins that were present across all TMT batches. Of these, 1,692 (91.4%) were quantified and underwent statistical analyses to identify proteomic differences across sample groups. According to the ANOVAs with Tukey posthoc test using MetaboAnalyst 5.0, 493 proteins were identified as differentially expressed proteins. Among the 279 proteins that changed with a statistically significant difference between D11 Exp and D11 Ctrl, 188 proteins were upregulated in D11 Exp including ACAT1, MDH2, SDHB, and ALDH2.

4.4.5 Molecular interaction network by injected methionine

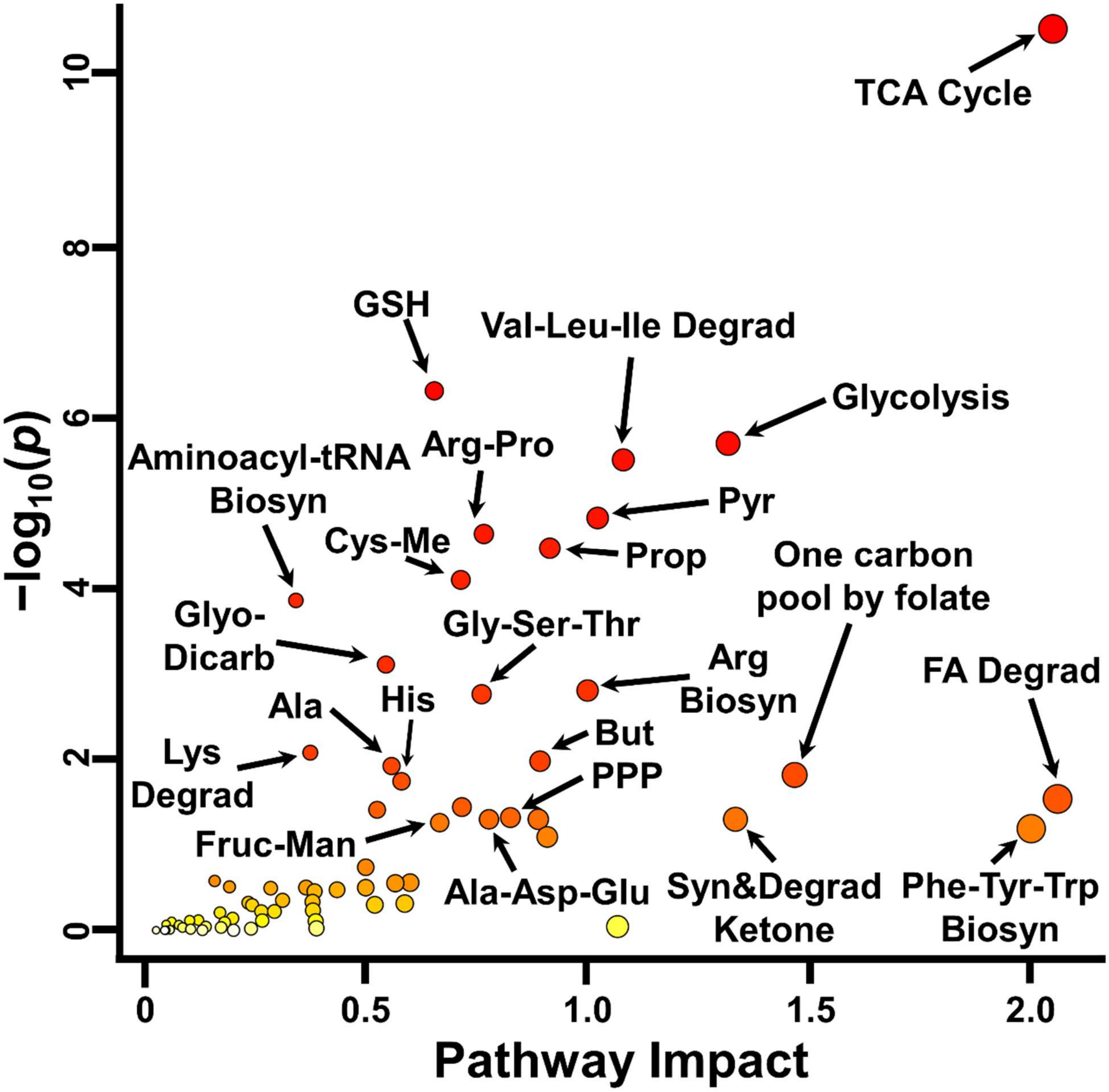


Figure 4.6 Revealed metabolic pathways by integrating identified metabolites and proteins using MetaboAnalyst 5.0. Joint pathway analysis: organism using *Homo sapiens* (human); pathway database using metabolic pathways (integrated). Algorithm selection: enrichment analysis (hypergeometric test); topology measure (degree centrality); integration method (combine queries). Notes: pathway abbreviation: pyruvate metabolism (Pyr), glutathione metabolism (GSH), arginine and proline metabolism (Arg-Pro), cysteine and methionine metabolism (Cys-Met), valine, leucine and isoleucine degradation (Val-Leu-Ile Degrad), pentose phosphate pathway (PPP), fatty acid degradation (FA Degrad), arginine biosynthesis (Arg Biosyn.), butanoate metabolism (But), histidine metabolism (His), fructose and mannose metabolism (Fruc-Man), alanine, aspartate, and glutamate metabolism (Ala-Asp-Glu), propanoate metabolism (Prop), lysine degradation (Lys Degrad), synthesis and degradation of ketone bodies (Syn&Degrad Ketone), phenylalanine, tyrosine and tryptophan biosynthesis (Phe-Tyr-Trp Biosyn), glycine, serine and threonine metabolism (Gly-Ser-Thr), glyoxylate and dicarboxylate metabolism (Glyo-Dicarb), beta-Alanine metabolism (Ala).

To uncover the perturbed metabolic pathways by injected methionine, we mapped the detected metabolites and proteins to known metabolic pathways on the KEGG database using MetaboAnalyst 5.0[188] (see details in **Table. 4.1**). **Fig. 4.6** shows the revealed metabolic pathways including TCA cycle (pathway impact: 2.05), GSH metabolism (pathway impact: 0.65), glycolysis (pathway impact: 1.32), pyruvate metabolism (pathway impact: 1.02), arginine-proline metabolism (pathway impact: 0.77), and cysteine-methionine metabolism (pathway impact: 0.71). Integrating metabolomic and proteomic datasets provided complementary pathway analysis results by revealing pathways that would be missed if only one 'omics was performed. Moreover, the multi-omic strategy expanded metabolic coverage for pathways with detected metabolites and proteins as shown by the example of the arginine-proline pathway.

By integrating these pathways in **Fig. 4.6** with metabolic and proteomic changes caused by injected methionine, we revealed perturbed metabolic pathways including arginine-proline metabolism and glycerophospholipid metabolism. After a literature review, we focused on the arginine-proline metabolism pathway because this pathway showed high pathway impact and interacted with the cysteine-methionine metabolism pathway through S-adenosylmethionine (SAM)[26, 165]. By scrutinizing the arginine-proline metabolism pathway, we found that 10 out of 38 metabolites were detected by HILIC-HRMS, and these metabolites showed either no significant change or downregulated change (p-value <0.05) in the D11 Exp compared to D11 Ctrl. Thus, we speculate that injected methionine caused the downregulation of the arginine-proline pathway in neurula stage 13 D11 cell clones. Furthermore, arginine-proline metabolism was found to be connected to glycine-serine-threonine metabolism by 4-hydroxyproline and pyruvate[165], which could explain the downregulation of serine and threonine caused by injected methionine. Glycine-serine-threonine interacted with cysteine-methionine pathway via threonine and interacted with arginine-proline pathway via serine.[165] These pathways interact with energy production pathways including glycolysis, PPP and fatty acids degradation via TCA cycle. Therefore, the detected metabolites formed a molecular interaction network as shown in **Fig 4.7**.

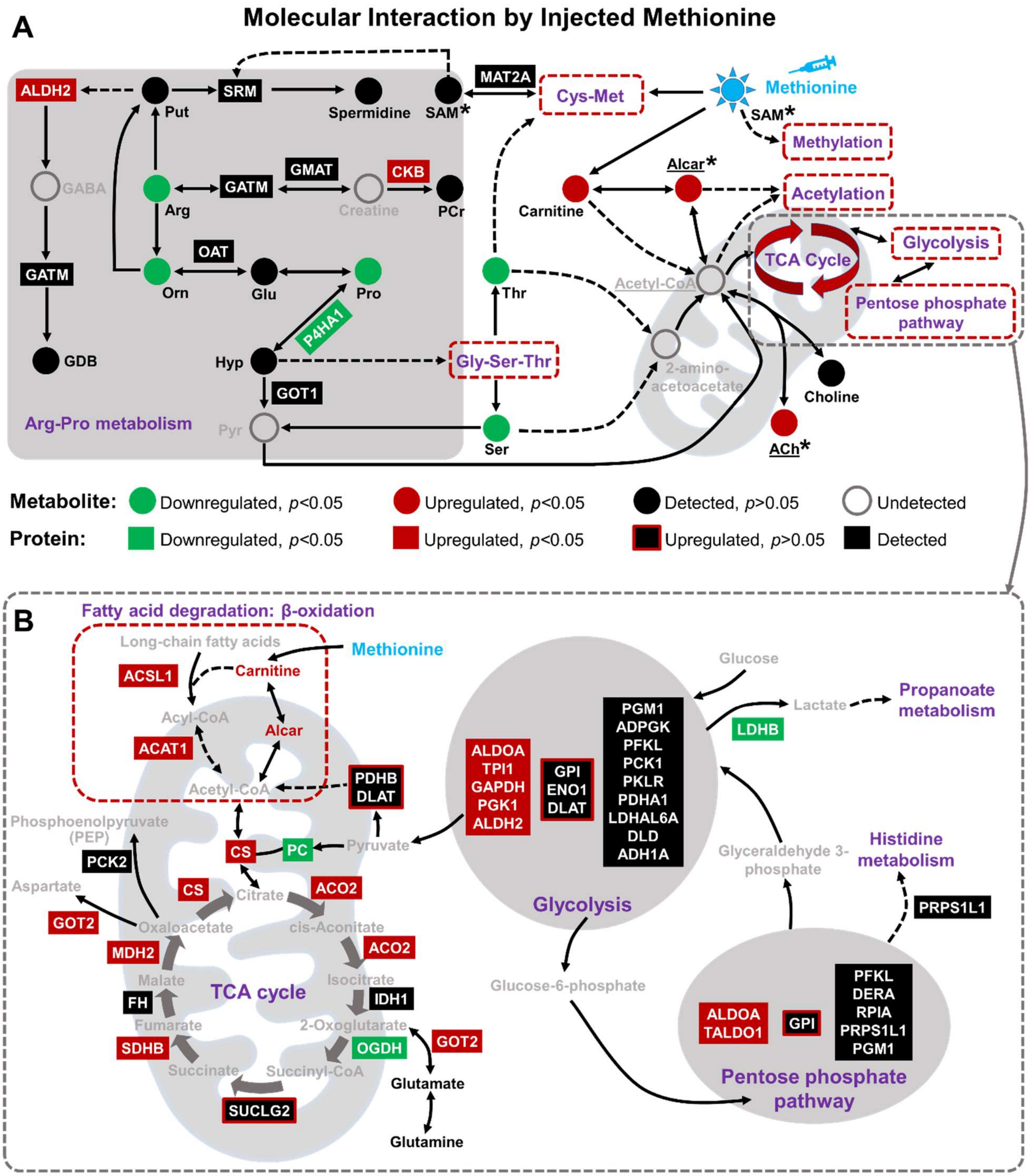


Figure 4.7 Molecular interaction network by injected methionine. (**A**) Revealed metabolic pathways using MetaboAnalyst 5.0 by mapping identified metabolites to the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. Metabolites with an asterisk (*) are metabolites incorporated with heavy isotope labeling from microinjected heavy methionine. Metabolites with an underscore (_) are metabolites that were validated to cause cell fate change of D11 from neural to epidermal tissues. (**B**) Zoomed in molecular interaction network involving pathways related to energy production.

Intriguingly, serine and threonine are both involved in the mitochondrial activities via acetyl coenzyme A (acetyl-CoA) according to a review paper by Chandel *et al.*[189] Specifically, threonine can be converted to 2-amino-3-oxobutanoic acid through a reversible reaction: threonine + $NAD^+ \rightleftharpoons$ 2-amino-3-oxobutanoic acid + NADH + H^+ in mitochondria.[165, 189] In turn, 2-amino-3-oxobutanoic acid can be converted to acetyl-CoA via a reversible reaction: 2-amino-3-oxobutanoic acid + CoA $\rightleftharpoons$ acetyl-CoA + glycine.[165, 189] Serine interacts with acetyl-CoA through the central carbon metabolism.[165, 189] Thus, we inferred that the downregulation of threonine and serine could affect the intracellular acetyl-CoA level. Acetyl-CoA is one of the mitochondria intermediates involved in energy production as well as protein PTMs. The active role that threonine and serine played in regulating cell fate decisions was revealed previously.[6, 190] Microinjection of threonine and serine, enriched in neural-fated D11 precursor cells, resulted in the epidermal-fated V11 cells changing cell fate from epidermal to neural tissue.[6] This indicated an increased level of threonine and serine caused the tissue fate to shift towards neural tissue while a lower level of threonine and serine maintained cells' epidermal tissue fate. Therefore, the observed (**Fig. 4.5**) downregulated level of threonine and serine was inferred to correlate to D11 cells' tissue fate shifting towards epidermal tissue.

While injected methionine can perturb the level of acetyl-CoA by downregulating serine and threonine, changing the concentration of methionine can alternatively perturb the intracellular acetyl-CoA level via carnitine and acetylcarnitine. Our metabolomic data showed that microinjection of methionine into the neural-fated D11 cells increased the abundance of carnitine in D11 Exp (**Fig. 4.5**) because methionine is a precursor of carnitine in its biosynthetic pathway[191-193]. Carnitine plays a crucial role in transporting long-chain fatty acids into mitochondria, the cell's power station for energy production,[168] for β-oxidation to produce acetyl-CoA.[190, 194] As both acetyl-CoA and carnitine can be converted into acetylcarnitine during this process[194], we expected to see an increased level of acetylcarnitine induced by injected methionine, which was supported by our experimental data showing enriched acetylcarnitine in D11 Exp compared to D11 Ctrl (**Fig. 4.5**). Moreover, acetylcholine, which could be reversibly converted from acetyl-CoA,[195] was also found to be enriched by injected methionine (**Fig. 4.5**). Therefore, when considering all the above clues, acetyl-CoA was in a hub that connected to the downregulated threonine and serine and upregulated carnitine, acetylcholine, and acetylcarnitine and therefore hypothesized to be one of the key molecular effectors of the methionine-induced cell fate change action (**Fig. 4.7A**).

4.4.6 Stable isotope labeling experiments to track the metabolic fate of methionine

To validate whether the injected methionine was converted to upregulated metabolites including carnitine, acetylcarnitine, and acetylcholine, we performed stable isotope labeling experiments to track the metabolic fate of methionine. D11 cell clones at neurula stage 13, originating from D11 cells injected with heavy methionine at the 16-cell stage, were manually dissected. By analyzing metabolite extracts from dissected D11 cell clones using a targeted HILIC-HRMS experiment, we found that acetylcholine and acetylcarnitine were incorporated with heavy isotope labeling from heavy methionine, indicating that injected methionine was finally converted to acetylcarnitine and acetylcholine as D11 cells differentiated during the cell fate change actions. We also found that S-adenosyl methionine (SAM), closely related to methionine metabolism and a direct product from methionine,[190] was incorporated with heavy isotope labeling. It is noteworthy that SAM showed no significant change by injecting light methionine while SAM incorporated the heavy isotope from heavy methionine. This could be because the injected methionine was converted to SAM,[196] which was then rapidly degraded or consumed in >100 reactions[196] by the time the embryos reach the neurula stage (~12 hours after injection). We expanded heavy isotope labeling analysis to all other metabolites detected in the molecular interaction network in **Fig. 4.7**. We found that all metabolites including arginine, proline, threonine, and serine showed no incorporation of heavy isotope labeling. This was surprising because no incorporation meant that the injected methionine was not directly converted to these metabolites involved in metabolic pathways including the cysteine-methionine pathway, arginine-proline pathway, and glycine-serine-threonine pathway. Thus, though these metabolic pathways were perturbed, their involvement in cell fate change actions could not be confirmed. The findings that acetylcarnitine, acetylcholine, and SAM incorporated heavy isotope labeling helped us to hypothesize potential molecular effectors including acetylcarnitine, acetylcholine, and acetyl-CoA involved in cell fate change actions.

4.4.7 Functional studies: D11 cell fate vs acetylcarnitine, acetyl-CoA, and serine/threonine

To validate that the hypothesized molecular effectors drive cell fate changes, functional studies were performed. The role of acetyl-CoA in regulating cell fate during cell differentiation was confirmed by performing functional studies. We performed the Mann-Whitney test for image scoring results between the wild-type larvae (N=20) and larvae from acetyl-CoA injected embryos (N=26) and found the injection of acetyl-CoA did cause the cell fate change of D11 cells from neural to epidermal tissues (**Fig. 4.8A and 4.8B**). Specifically, the scoring of neural tissues including the retina, brain, central somite, and olfactory decreased significantly with *p*-values of 2.0×10^{-7}, 6.2×10^{-4}, 3.6×10^{-3}, and 8.9×10^{-5}, respectively. In contrast, the scoring of head epidermal and body epidermal tissues increased significantly (*p*-value of 9.2×10^{-6}, and 3.9×10^{-3}, respectively).

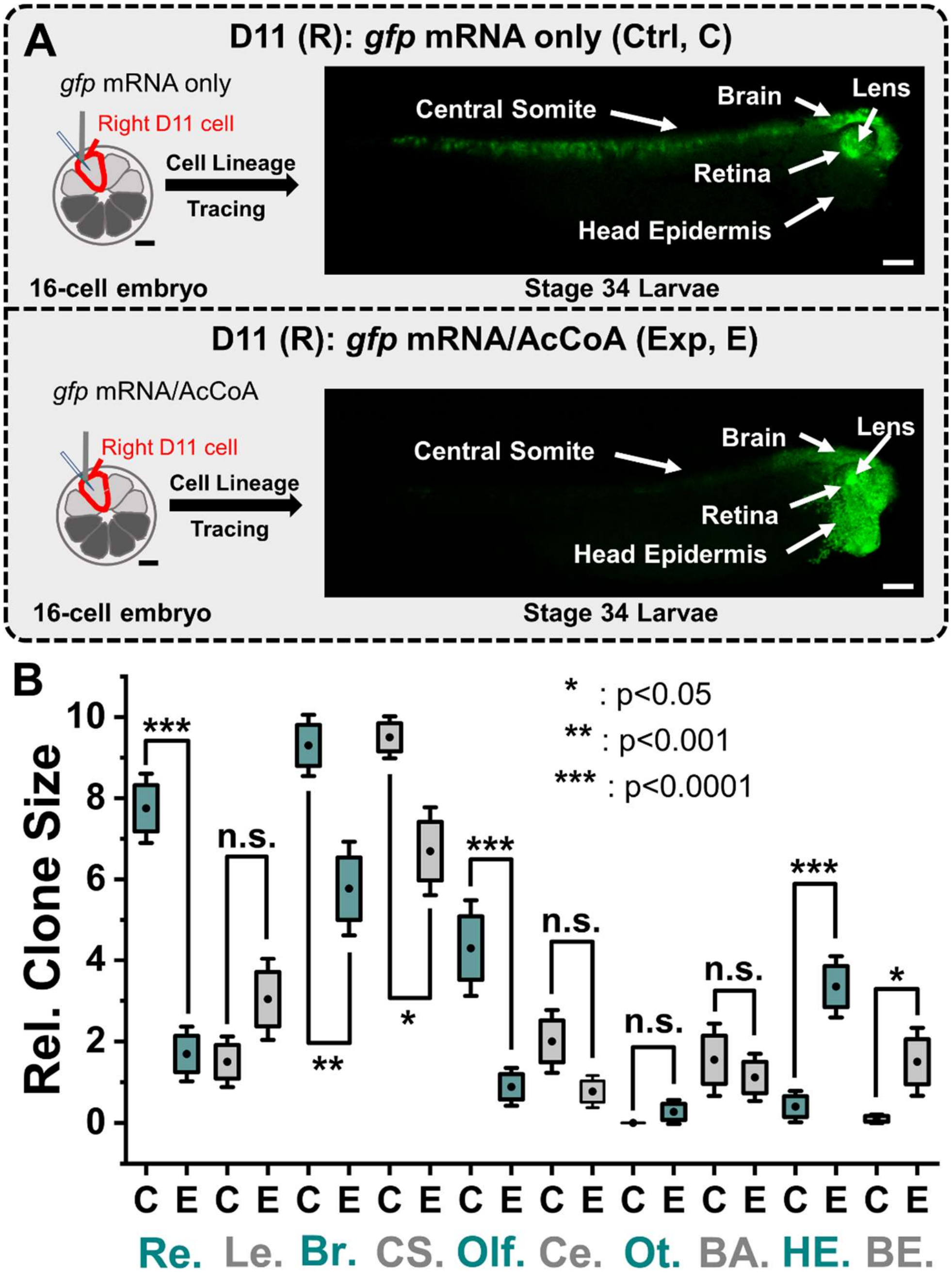


Figure 4.8 Functional studies of acetyl-CoA as the potential effector to cause cell fate change. The workflow included microinjection, cell lineage tracing, embryonic fixation, and image scoring. (**A**) Functional study of acetyl-CoA. Top panel: right D11 cells in 16-cell embryos were microinjected with *gfp* mRNA and injected embryos were cultured to larval stage 34 as the control group (Ctrl, C, N=20); Bottom panel: right D11 cells in 16-cell embryos were microinjected with a mixture of *gfp* mRNA and acetyl-CoA and injected embryos were cultured to larval stage 34 as the experimental group (Exp, E, N=26). Scale bar: 250 µm. (**B**) Statistical analysis of tissue image scoring between Ctrl and Exp for functional study of acetyl-CoA. Box-whisker plots: Box with 1×standard error of the mean (SEM) and whisker with 1.5×SEM. Key: n.s., non-significant, *$p<0.05$, **$p<0.001$, ***$p<0.0001$, Mann-Whitney test. Box-whisker plots: Box with 1×standard error of the mean (SEM) and whisker with 1.5×SEM. Key: n.s., non-significant, *$p<0.05$, **$p<0.001$, ***$p<0.0001$, Mann-Whitney test. Scored tissue name abbreviation: Retina (Re.); Lens (Le.); Brain (Br.); Central somite (CS.); Olfactory placode (Olf.); Cement gland (Ce.); Otocyst (Ot.); Branchial arches (BA.); Head epidermis (HE.); Body epidermis (BE.).

Similarly, acetylcarnitine was microinjected to right D11 cells into 16-cell embryos and the injected embryos were cultured to larva stage 34. The fixed larvae were imaged and scored following a previously established protocol.[6] Mann-Whitney test was performed between the wild-type larvae (N=14) and larvae from acetylcarnitine-injected embryos (N=17). We found that acetylcarnitine did cause the D11 cells to change cell fate from neural to epidermal tissues. Specifically, the scoring of neural tissues including the retina, brain, central somite, and olfactory placode decreased significantly with p-values of 5.1×10^{-5}, 1.8×10^{-5}, 4.7×10^{-4}, and 1.3×10^{-4}, respectively. In contrast, the scoring of head epidermal tissues increased significantly (p-value of 5.5×10^{-6}) (**Fig. 4.9A)**. Moreover, a previous study found that injection of acetylcholine alone into D11 cells at the 16-cell stage also caused the cell fate change from neural to epidermal tissue.[59]

In contrast, we performed the functional study for the mixture of serine and threonine, which served as a negative control because they both were downregulated by injected methionine and showed no incorporation of heavy isotope labeling from methionine. We predicted that the injected serine and threonine should cause no cell fate change of D11 cells, which was confirmed by the Mann-Whitney test between the wild-type larvae (N=18) and the experimental group (N=16), showing the injected serine/threonine did not cause significant change (p-value <0.05) for neural tissues including the retina, brain, central somite, and olfactory placode (**Fig. 4.9B)**. Therefore, serine and threonine are not molecular effectors that are involved in methionine-induced cell fate change. Moreover, we found that injection of serine/threonine mixture did promote the formation of the lens from D11 cells (p-value of 0.007). This could be explained by the fact that D11 cells are precursor cells of the lens along with V11 cells,[72] which were also revealed to form more lens tissue by the injected serine/threonine mixture.[6]

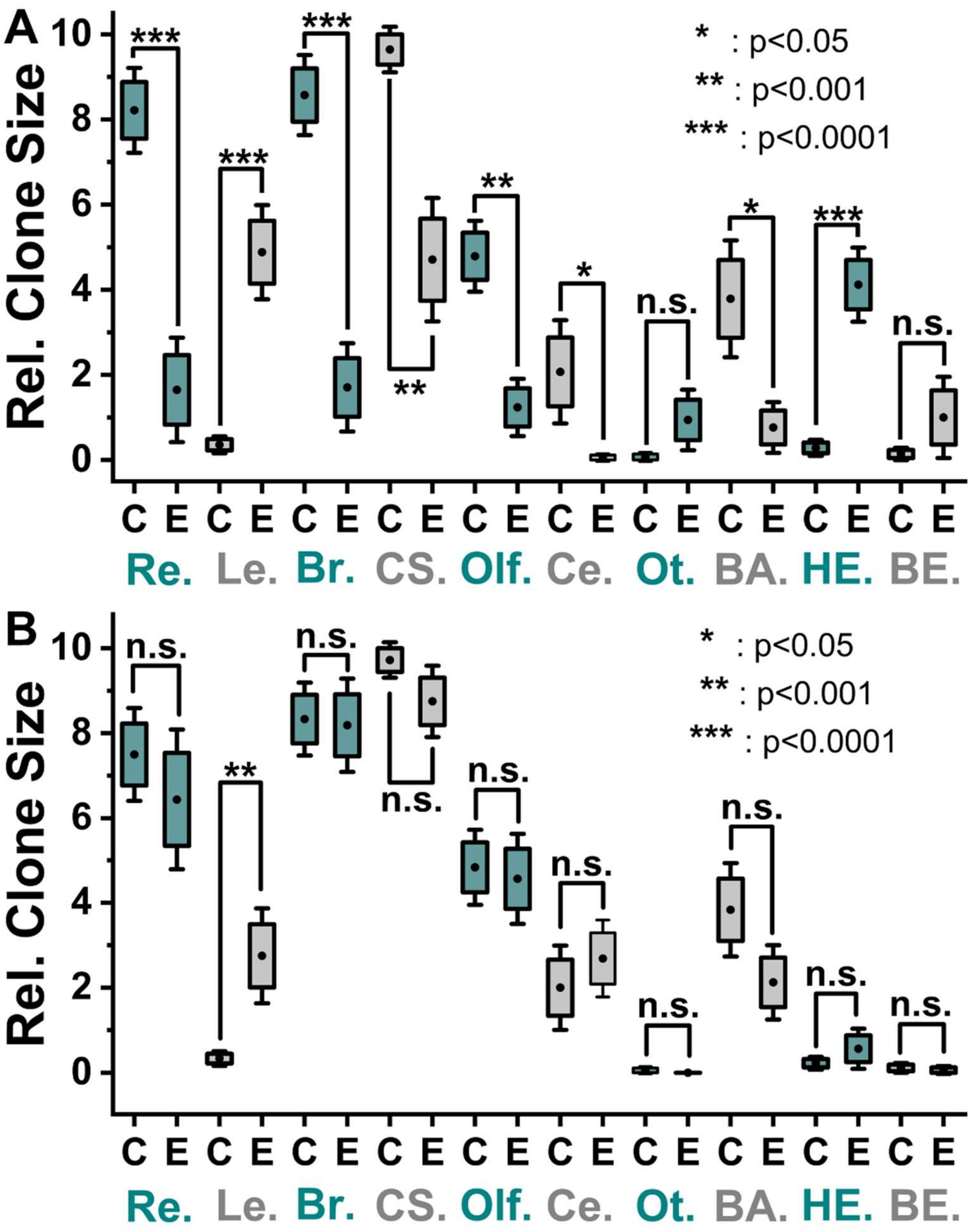


Figure 4.9 Functional studies of acetylcarnitine and serine/threonine to cause cell fate change. The workflow included microinjection, cell lineage tracing, embryonic fixation, and image scoring. (**A**) Statistical analysis of tissue image scoring between Ctrl (N=14) and Exp (N=17) for functional study of acetylcarnitine. Box-whisker plots: Box with 1×standard error of the mean (SEM) and whisker with 1.5×SEM. Key: n.s., non-significant, *$p<0.05$, **$p<0.001$, ***$p<0.0001$, Mann-Whitney test. (**B**) Statistical analysis of tissue image scoring between Ctrl (N=18) and Exp (N=16) for functional study of serine/threonine. Box-whisker plots: Box with 1×standard error of the mean (SEM) and whisker with 1.5×SEM. Key: n.s., non-significant, *$p<0.05$, **$p<0.001$, ***$p<0.0001$, Mann-Whitney test. Scored tissue name abbreviation: Retina (Re.); Lens (Le.); Brain (Br.); Central somite (CS.); Olfactory placode (Olf.); Cement gland (Ce.); Otocyst (Ot.); Branchial arches (BA.); Head epidermis (HE.); Body epidermis (BE.).

4.4.8 Proposed mechanism underlying methionine-induced cell fate change: perturbation of acetylation PTMs

The functional validation experiments supported that the injected acetylcarnitine, acetylcholine, and acetyl-CoA altered the cell fate change of neural precursor D11 cells, which enabled us to generate hypotheses of the mechanisms underlying the cell fate change actions along with the metabolomic/proteomic results and stable isotope labeling experiments. We proposed that injected methionine perturbed the intracellular acetyl-CoA level, which affected gene expression by facilitating protein acetylation PTMs and perturbing energy production pathways.

Acetyl-CoA provides acetyl groups for protein acetylation, which is a major PTM of proteins[197-198] For example, acetyl-CoA is a crucial player in histone protein acetylation[190, 199] by working as a substrate of an enzyme that catalyzes the acetylation reaction.[200] The level of histone acetylation was revealed to increase with the supply increase of acetyl-CoA and can be regulated by varying the level of acetyl-CoA.[199] Histone acetylation was known to cause the change of chromatin structure and thus involved in activating transcription by affecting gene expression.[190] Additionally, acetyl-CoA was found to regulate the cell differentiation of human and mouse embryonic stem cells via histone acetylation.[201] Furthermore, acetyl-CoA provides the acetyl group to synthesize UDP-N-acetylglucosamine (UDP-GlcNAc), which caused the O-linked N-acetylglucosamine (O-GlcNAc) PTM to the serine/threonine residues of target proteins,[202] thus, affecting cellular activities such as transcription and signaling.[203] We found in our proteomic results that histone H4 proteins, a revealed O-GlcNAc site,[203] were upregulated by injected methionine, indicating a higher abundance of histone acetylation, which, in turn, affects cellular processes.

We further checked the quantified acetylated proteins to explore potential acetylation changes by the injected methionine. Since the TMT-labeling strategy used in the proteomic analysis is not intended for detecting acetylated proteins, the number of detected acetylated proteins in this study was limited. However, 9 out of 10 detected acetylated proteins showed upregulation by injected methionine and 2 out of the 9 were upregulated with statistically significant differences ($p<0.05$, one-way ANOVAs with Tukey's posthoc test). Therefore, these results showed a sign of the upregulation of acetylation PTMs by injected methionine.

To validate that acetylation PTM can cause cell fate change, I performed a functional study using sodium butyrate which is a short-chain fatty acid that was previously proved to inhibit the activity of histone deacetylase to promote histone acetylation for various types of cells[204-205]. According to the Mann-Whitney test for scoring results between the wild-type larvae (N=20) and larvae from sodium butyrate injected embryos (N=22), the injected sodium butyrate did cause the cell fate change of D11 cells from neural to epidermal tissues (**Figure 4.10A and 4.10B**). Specifically, the size of neural tissues including the retina, brain, central somite, and olfactory decreased significantly with *p*-values of 1.9×10^{-6}, 1.8×10^{-4}, 1.1×10^{-4}, and 8.4×10^{-4}, respectively. In contrast, the scoring of head epidermal and body epidermal tissues increased significantly (*p*-value of 3.3×10^{-6}, and 1.0×10^{-2}, respectively). Therefore, the functional study results of sodium butyrate demonstrated that upregulated acetylation of histone did induce D11 cells to change cell fate from neural to epidermal tissues, which in turn supported the hypothesis that upregulated acetylation PTM induced by methionine is responsible for the cell fate change actions.

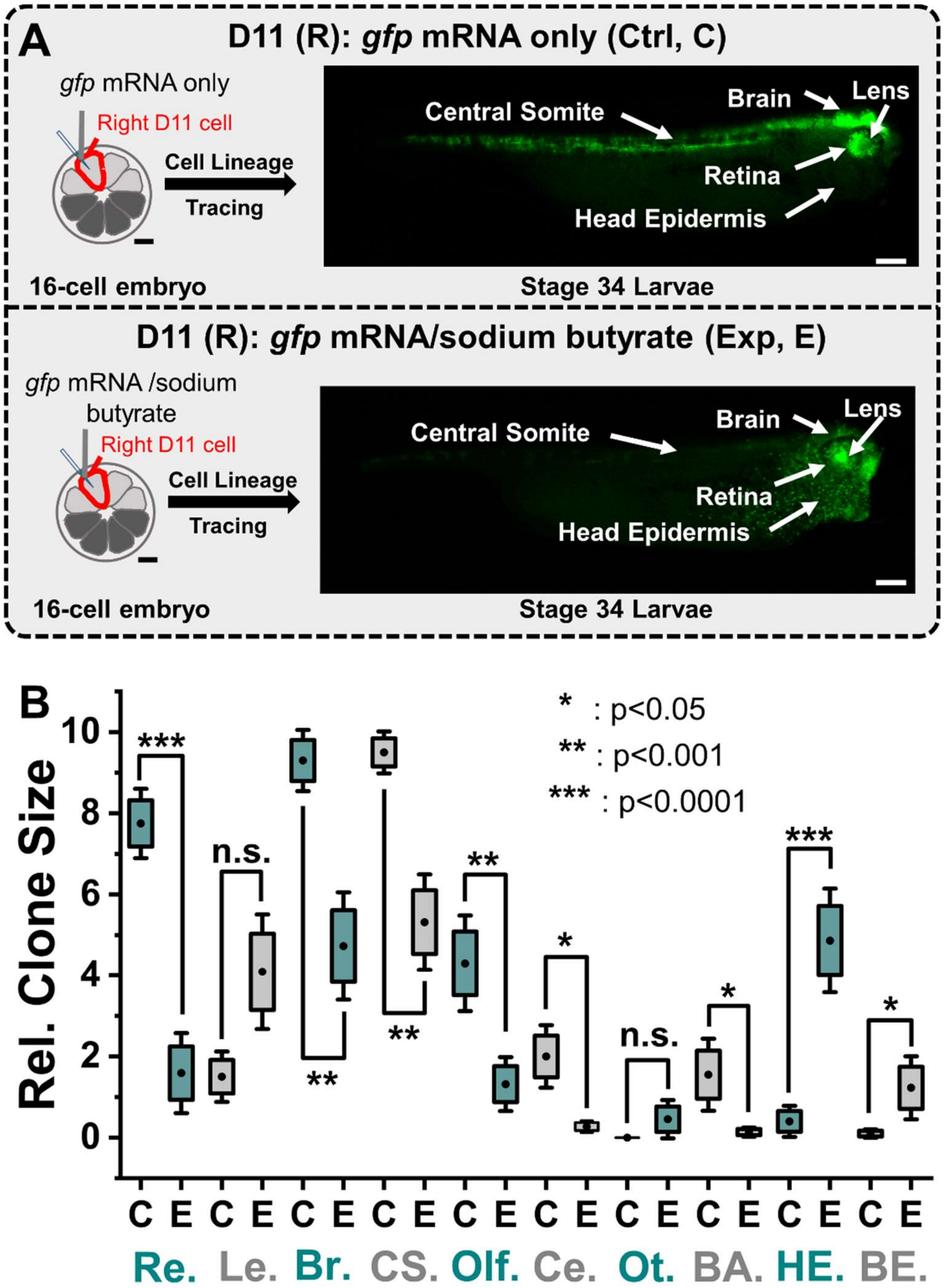


Figure 4.10 Functional study to test the role of acetylation in cell fate change. The workflow included microinjection, cell lineage tracing, embryonic fixation, and image scoring. (**A**) Functional study of sodium butyrate. Top panel: right D11 cells in 16-cell embryos were microinjected with *gfp* mRNA and injected embryos were cultured to larval stage 34 as the control group (Ctrl, C, N=20); Bottom panel: right D11 cells in 16-cell embryos were microinjected with a mixture of *gfp* mRNA and sodium butyrate and injected embryos were cultured to larval stage 34 as the experimental group (Exp, E, N=22). Scale bar: 250 µm. (**B**) Statistical analysis of tissue image scoring between Ctrl and Exp for functional study of sodium butyrate. Box-whisker plots: Box with 1×standard error of the mean (SEM) and whisker with 1.5×SEM. Key: n.s., non-significant, *$p<0.05$, **$p<0.001$, ***$p<0.0001$, Mann-Whitney test. Scored tissue name abbreviation: Retina (Re.); Lens (Le.); Brain (Br.); Central somite (CS.); Olfactory placode (Olf.); Cement gland (Ce.); Otocyst (Ot.); Branchial arches (BA.); Head epidermis (HE.); Body epidermis (BE.).

4.4.9 Proposed mechanism underlying methionine-induced cell fate change: perturbation of cellular energy production

Although evidence corroborated that acetylation PTM was involved in methionine-induced cell fate change actions, injected methionine could also perturb energy production to cause cell fate change. Carnitine and acetylcarnitine involve in transferring long-chain fatty acids into mitochondria for β-oxidation to produce ATP and acetyl-CoA. The generated acetyl-CoA is further involved in mitochondria energy production as a key metabolite of the TCA cycle.[189, 194, 197] Therefore, upregulated carnitine and acetylcarnitine induced by injected methionine facilitated the production of energy molecule ATP and acetyl-CoA. The direct evidence is that both long-chain fatty-acid-coenzyme A ligase 1 (ACSL1) and acetyl-CoA acetyltransferase 1 (ACAT1), playing important roles in fatty acid degradation, were upregulated by injected methionine according to the proteomic results. ACSL1 involves in converting long-chain fatty acids to acyl-CoAs[206] while ACAT1 catalyzes the reversible conversion from acetoacetyl-CoA to acetyl-CoA[207] (**Fig. 4.7B**). Since acetyl-CoA fuels citrate in the TCA cycle for energy production in mitochondria, the increased acetyl-CoA level from upregulated carnitine and acetylcarnitine was further supported by our proteomic results showing the upregulation of the TCA cycle (**Fig. 4.7B**). Specifically, 5 out of the detected 10 proteins involved in TCA cycle[208] including CS, ACO2, MDH2, SDHB, and GOT2 were upregulated with statistical significance ($p<0.05$) by injected methionine while OGDH was downregulated. The rest of the 3 proteins were either upregulated with no significant change ($p>0.05$) or unchanged. The upregulation of the TCA cycle supported the hypothesis that the injected methionine perturbed energy production.

Likewise, we examined another energy production pathway called glycolysis which converts glucose to pyruvate and interacts with the TCA cycle via pyruvate. The proteomic results indicated the upregulation of glycolysis by injected methionine (**Fig. 4.7B)**. Particularly, 8 proteins were upregulated from the 17 detected proteins in the glycolysis pathway, among which 5 proteins (ALDOA, TPI1, GAPDH, PGK1, and ALDH2) showed a significant difference ($p<0.05$). Moreover, the pentose phosphate pathway (PPP) is another pathway related to energy production and is closely connected to glycolysis by providing glyceraldehyde 3-phosphate to glycolysis and obtaining glucose-6-phosphate from glycolysis. Investigating the detected proteins involved in PPP uncovered these proteins were either upregulated or remained unchanged by injected methionine, suggesting the PPP was upregulated (**Fig. 4.7B**). The revealed upregulation of the TCA cycle, glycolysis, and PPP represented an upregulated cellular energy production, suggesting a hypothesis of cell fate change caused by cellular energy changes by injected methionine as cellular energy production was revealed to instruct cell fate decision according to various literature reviews[168, 209].

4.5 Conclusions

Technological advances have deepened our understanding of the roles of biomolecules in biological activities. Metabolites were revealed to be not only passive players but also active drivers of various biological processes such as cell differentiation, aging, and disease development. To understand the role of metabolites in driving cell differentiation, we explored the mechanisms underlying metabolite-induced cell fate changes by combining embryonic dissociation and FACS to isolate tissue-specific cell clones for LC-MS metabolomics and proteomics. The multi-omic strategy empowered us to reveal perturbed metabolic pathways and to identify potential molecular effectors responsible for cell fate change actions. The obtained results enabled us to generate hypotheses that were partially validated using multi-omic data, stable isotope labeling, and functional study, promising to help unravel molecular mechanisms underlying metabolite-induced cell fate changes.

Table 4.1 Metabolic pathways revealed by joint pathway analysis using MetaboAnalyst 5.0. FDR: false discovery rate.

Pathway name	Total	Hits	Raw *p*	-log10(*p*)	FDR	Impact
Citrate cycle (TCA cycle)	42	20	2.56E-11	10.592	2.15E-09	2.05
Glutathione metabolism	56	18	4.55E-07	6.342	1.91E-05	0.65
Glycolysis or Gluconeogenesis	61	18	1.89E-06	5.724	5.29E-05	1.32
Valine, leucine and isoleucine degradation	88	22	2.94E-06	5.531	6.18E-05	1.08
Pyruvate metabolism	45	14	1.42E-05	4.849	2.38E-04	1.02
Arginine and proline metabolism	78	19	2.18E-05	4.661	3.06E-04	0.77
Propanoate metabolism	48	14	3.22E-05	4.492	3.87E-04	0.91
Cysteine and methionine metabolism	71	17	7.65E-05	4.117	8.03E-04	0.71
Aminoacyl-tRNA biosynthesis	74	17	1.33E-04	3.875	1.24E-03	0.34
Glyoxylate and dicarboxylate metabolism	56	13	7.55E-04	3.122	6.35E-03	0.55
Arginine biosynthesis	27	8	1.53E-03	2.817	1.16E-02	1.00
Glycine, serine and threonine metabolism	68	14	1.69E-03	2.773	1.18E-02	0.76
Lysine degradation	49	10	8.19E-03	2.087	0.0529	0.38
Butanoate metabolism	29	7	0.010	1.986	0.0619	0.89
beta-Alanine metabolism	44	9	0.012	1.930	0.0659	0.56
One carbon pool by folate	31	7	0.015	1.825	0.0786	1.47
Histidine metabolism	32	7	0.018	1.750	0.0879	0.58
Fatty acid degradation	102	15	0.029	1.543	0.1336	2.06
Alanine, aspartate and glutamate metabolism	61	10	0.036	1.447	0.1578	0.72
Amino sugar and nucleotide sugar metabolism	79	12	0.038	1.416	0.1613	0.53
Pentose phosphate pathway	47	8	0.047	1.324	0.1739	0.83
Nitrogen metabolism	10	3	0.050	1.304	0.1739	0.78
D-Glutamine and D-glutamate metabolism	10	3	0.050	1.304	0.1739	0.89
Synthesis and degradation of ketone bodies	10	3	0.050	1.304	0.1739	1.33
Fructose and mannose metabolism	40	7	0.054	1.264	0.1828	0.67
Phenylalanine, tyrosine and tryptophan biosynthesis	11	3	0.064	1.194	0.2069	2.00
Valine, leucine and isoleucine biosynthesis	12	3	0.080	1.097	0.2491	0.91

Riboflavin metabolism	9	2	0.182	0.740	0.5458	0.50
Drug metabolism - other enzymes	70	8	0.263	0.579	0.7592	0.16
Phenylalanine metabolism	21	3	0.276	0.560	0.7592	0.60
Folate biosynthesis	61	7	0.280	0.553	0.7592	0.57
Tryptophan metabolism	84	9	0.309	0.510	0.7685	0.19
Porphyrin and chlorophyll metabolism	53	6	0.315	0.502	0.7685	0.37
Ascorbate and aldarate metabolism	13	2	0.316	0.501	0.7685	0.50
Starch and sucrose metabolism	43	5	0.320	0.495	0.7685	0.29
Glycerophospholipid metabolism	86	9	0.334	0.476	0.7800	0.44
Thiamine metabolism	14	2	0.349	0.457	0.7924	0.38
Linoleic acid metabolism	17	2	0.445	0.352	0.9831	0.31
Galactose metabolism	51	5	0.463	0.334	0.9891	0.38
Sulfur metabolism	18	2	0.475	0.323	0.9891	0.24
Purine metabolism	166	15	0.485	0.314	0.9891	0.59
Pyrimidine metabolism	99	9	0.502	0.299	0.9891	0.52
Nicotinate and nicotinamide metabolism	42	4	0.506	0.296	0.9891	0.24
alpha-Linolenic acid metabolism	22	2	0.585	0.233	1.0000	0.38
Selenocompound metabolism	35	3	0.601	0.221	1.0000	0.26
Glycerolipid metabolism	35	3	0.601	0.221	1.0000	0.29
Terpenoid backbone biosynthesis	36	3	0.621	0.207	1.0000	0.17
Arachidonic acid metabolism	81	6	0.724	0.140	1.0000	0.20
Sphingolipid metabolism	58	4	0.760	0.119	1.0000	0.12
Drug metabolism - cytochrome P450	98	7	0.767	0.115	1.0000	0.10
Glycosphingolipid biosynthesis-globo and isoglobo series	31	2	0.768	0.114	1.0000	0.27
Pentose and glucuronate interconversions	32	2	0.784	0.106	1.0000	0.39
Ubiquinone and other terpenoid-quinone biosynthesis	17	1	0.789	0.103	1.0000	0.06
Pantothenate and CoA biosynthesis	34	2	0.812	0.091	1.0000	0.18
Biotin metabolism	21	1	0.854	0.068	1.0000	0.05
Ether lipid metabolism	39	2	0.868	0.062	1.0000	0.08
Fatty acid elongation	75	4	0.905	0.043	1.0000	1.07

Glycosaminoglycan degradation	44	2	0.908	0.042	1.0000	0.14
Glycosphingolipid biosynthesis-ganglio series	47	2	0.927	0.033	1.0000	0.09
Retinol metabolism	47	2	0.927	0.033	1.0000	0.17
Metabolism of xenobiotics by cytochrome P450	145	8	0.949	0.023	1.0000	0.39
Tyrosine metabolism	88	4	0.957	0.019	1.0000	0.24
N-Glycan biosynthesis	77	3	0.971	0.013	1.0000	0.11
Inositol phosphate metabolism	69	2	0.987	0.006	1.0000	0.06
Primary bile acid biosynthesis	92	2	0.998	0.001	1.0000	0.13
Phosphatidylinositol signaling system	74	1	0.999	0.000	1.0000	0.03
Steroid biosynthesis	82	1	1.000	0.000	1.0000	0.05
Steroid hormone biosynthesis	199	5	1.000	0.000	1.0000	0.20
Fatty acid biosynthesis	129	1	1.000	0.000	1.0000	0.05

Chapter 5: MagCar: a novel technique for temporal study of tissue formation

Based on manuscript draft in preparation for submission by **Jie Li**, Kaitlyn E. Stepler, and Peter Nemes*

My contributions (the lead author): I designed the research; I developed and validated the MagCar technology; I collected samples, conducted metabolomic experiments and data processing, and performed functional study; I analyzed the data, interpreted results, and wrote the manuscript draft.

5.1 Abstract

Isolating tissue-specific cells from developing organisms is crucial for studying cell differentiation, understanding tissue formation, and discovering disease biomarkers. Developing sampling techniques to isolate tissue-specific cells amenable to metabolomic, proteomic, transcriptomic, and genomic analysis has been a major aim of life science research. Generally, developing organisms have a fast cell cleavage rate, rapidly changing cell size, a vast number of descendant cells, and dispersed location of tissue-specific cells due to cell migration. All these factors increase the complexity and challenges of isolating tissue-specific cells using existing techniques such as dissection, microsampling, and cell sorting. For example, it is difficult and laborious to dissect dispersed tissue-specific cells without contamination from neighboring cells. Therefore, to harvest tissue-specific cell clones from developing organisms, an efficient sampling technique should be specific to label and track the development of cell clones, inclusive to collect dispersed and enormous cell clones, and fast enough to isolate cell clones in

their native state. To address these challenges, we developed a technique called MagCar which incorporated intracellular microinjection of fluorescent magnetic particles, cell lineage tracing, embryonic dissociation, and magnetic cell isolation to harvest tissue-specific cells for mass spectrometry (MS)-based multi-omics with fast speed, high purity, and cell size scalability. The developed MagCar sampling technique enables to study tissue formation during early-stage embryonic cell differentiation.

5.2 Introduction

A myriad of sampling techniques has been developed to collect samples for cell and developmental biology. Manual surgical dissection guided by cell morphology was developed to isolate single cells. One example is that different subtypes of neurons with relatively large sizes were dissected from the central nervous system of *Aplysia californica* to study the metabolic differences and similarities.[42] Additionally, embryonic cells with different tissue fates were dissected according to stereotypical cleavage and pigmentation, which revealed cell heterogeneity of the metabolome and proteome among tissue-specific precursor cells[6, 30]. Manual dissection was further developed by combining with fluorescent cell lineage tracing to isolate tissue-specific cells at later stages from developing organisms for downstream ‘omic analysis.[22, 210-211]

Aspiration using needles or syringes is another widely used technique for the manual isolation of tissue-specific cells.[210-211] For example, a fine needle was used to aspirate cells from tumors for DNA analysis[212] and to isolate neuronal cells from cell culture for RNA amplification.[213] To minimize perturbation to the sampling cells, a microprobe sampling technique was developed to aspirate cellular contents from individual cells directly, which was extended to study the subcellular distribution of biomolecules in single neurons.[28]

Microsampling was also applied to uncover cell heterogeneity[26] and metabolic reorganization during the neural tissue-fated cell clones development[43]. Recently, microsampling was demonstrated to enable multi-omic analysis of early-stage embryonic cells without interfering with normal frog embryonic development.[26]

Manual dissection and aspiration are considered to be high purity while having relatively low throughput and relying on the dexterity of manual operation,[210] which face challenges to isolate enormous and dispersed tissue-specific cells from developing organisms. Laser capture microdissection (LCM) was proposed as an automated cell isolation technique that improves throughput and preserves the three-dimensional structure of isolated tissue-specific cells.[108, 210] In LCM, a laser combined with a visual microscope is used to automatically dissect out cells of interest directly or to dissect out undesired cells, which minimizes the contamination of neighboring cells.[214-215] Cells harvested by LCM enable the later molecular analysis such as discovery proteomics and genomic sequencing.[215] However, LCM is primarily applied to fixed/frozen tissues, which limits its application to studying live organisms.

Cell sorting techniques including fluorescence-activated cell sorting (FACS) and magnetic-activated cell sorting (MACS) provide an alternative way to isolate cells with specific markers from developing organisms. FACS is widely used as it allows selective isolation of tissue-specific cells labeled by a fluorescent dye with high purity, throughput, and viability[210, 216]. During FACS, cells of interest with a fluorescent label are isolated based on cell size, internal complexity, and fluorescence intensity.[171] However, FACS has an intrinsic limitation on workable cell size and requires a large starting number of cells.[217] Cells with large sizes cause clogging issues to the flow chamber and orifice,[218] which results in the loss of samples and extra work for the maintenance of the sophisticated FACS instrument[218]. Moreover, cells are encapsulated by sheath flow to form single droplets and sorted by applying electrostatic potential, which may result in morphological and molecular changes to sorted cells.[218-219] Recent examples demonstrated that FACS caused oxidative stress and metabolic changes in sorted cells.[219-220]

MACS is another sorting technique, in which cells that have intrinsic or extrinsic magnetic characteristics are manipulated and sorted out based on magnetic force.[170, 221] For example, deoxyhemoglobin red blood cells containing native magnetic properties were isolated from whole blood under an applied magnetic field.[222] To isolate cells without intrinsic magnetic characteristics, magnetic particles are used in conjugation with antibodies that recognize cell-type-specific proteins to bind the surface of targeted cells, followed by magnetic cell sorting to isolate these targeted cells.[216, 221] Therefore, this sorting method is restricted to isolating tissue-specific cells with a known surface marker. For tissue-specific cells that are not yet known to have a surface marker, it is challenging to use MACS to selectively label and isolate these cells.

Though currently available techniques have had unprecedented benefits for the study of various complex questions, unfortunately, these available techniques are yet to be improved to isolate tissue-specific cell clones with high purity, high throughput, minimal perturbation to cells, scalability to wide cell size range, and minimal sample loss. For example, the gastrula stage of *X. laevis* embryos, a stage at which cells migrate to establish three different layers towards different tissues and organs, is a critical stage to understanding cell differentiation and organogenesis. However, there are no effective techniques to isolate tissue-specific cell clones from this stage because the cell clones have thousands of descendant cells and thus are tedious and difficult to isolate using existing techniques such as microsampling and microdissection. Furthermore, cells in the gastrula stage or earlier stages are too large (ranging from 40–200 μm in diameter) to be sorted using FACS. Since *X. laevis* cells are not yet known to express specific surface markers, it is difficult to apply the current MACS workflow to isolate tissue-specific cells. Thus, there are unmet but pressing needs to develop technology to isolate tissue-specific cells to study embryonic development.

Uncovering the temporal changes of biomolecules as tissue-fated precursor cells develop from single cells to tissue-specific cell clones and later respective tissues helps to understand regulation mechanisms underlying cell differentiation and tissue formation. These biomolecules, including upstream biomolecules such as RNA and DNA and downstream biomolecules such as proteins and metabolites, have distinct roles and interact through complex molecular networks. Revealing temporal dynamic changes of these different types of biomolecules during tissue formation provides insight into the underlying mechanisms involved in health and disease. Studies focused on either proteomic or metabolic changes during embryonic development and cell differentiation have been conducted. Metabolic and proteomic reorganizations were revealed at the single-cell level[43, 52] and whole-embryo level[7] during development. However, most of these previous studies only focused on one ‘ome—one type of biomolecule, which limits our complete understanding of the complex biological processes. Furthermore, though there were studies conducted using multi-omics, these studies were generally conducted by collecting different samples for each ‘omics analysis, thus precluding the application of this multi-omics strategy when samples are limited, and it is not possible to obtain multiple individual samples. Therefore, it is imperative to investigate temporal molecular changes via multi-omics of the same samples, which promises to provide a holistic view of a single precursor cell that differentiates toward pre-determined tissues.

To address the bottleneck to isolate tissue-specific cell clones from developing organisms to study tissue formation, we developed a sampling technique called cell magnetic cargo for cell sorting (MagCar), which (**Fig. 5.1**) incorporated intracellular microinjection of fluorescent magnetic particles (FMPs), cell lineage tracing, embryonic dissociation[79], and magnetic cell isolation. Unlike traditional approaches, we intracellularly injected FMPs into progenitor cells of specific tissue to label descendant cells. We designed a cell sorting apparatus and demonstrated its capability to isolate tissue-specific cells from different developmental stages of *X. laevis* embryos. We applied MagCar to study cell dorsal and ventral patterning to respective tissues. Specifically, we applied the developed workflow to isolate neural-fated and epidermal-fated cell clones from five key developmental stages covering blastula, gastrula, and neurula-stage *X. laevis* embryos. For the MagCar isolated cell clones, we further proved that the isolated cell clones enabled consecutive isolation of metabolites and proteins from the same sample. Thus, we achieved multi-omics to investigate the dynamic molecular changes during neural and epidermal tissue formation by mass spectrometry (MS)-based metabolomics and proteomics. This demonstrated MagCar's adaptability to other technologies and workflows for downstream analysis. Though we used *X. laevis* as our research model, the cell clone isolation methods and protocols developed in this study can be extended to other important biological models, such as zebrafish, which has an established cell fate map of early-stage progenitors[223], to elucidate how cells develop into tissues, organs, and whole organisms during normal and impaired development.

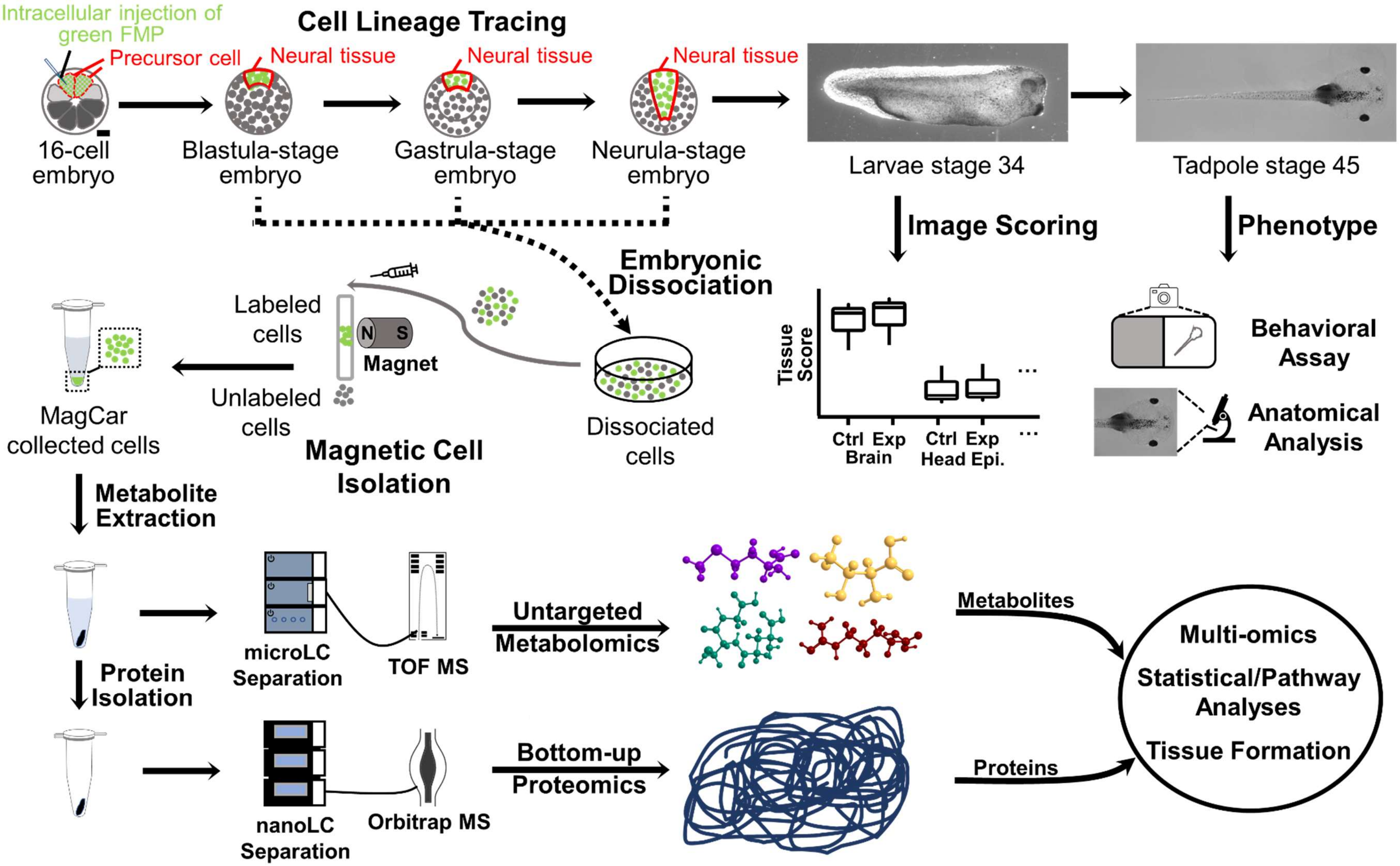


Figure 5.1 Workflow to isolate tissue-specific cell clones using MagCar from *Xenopus* frog embryo for metabolomic and proteomic analysis. This workflow integrates intracellular microinjection of fluorescent magnetic particles, cell lineage tracing, embryo dissociation, magnetic cell separation, simultaneous metabolite and protein isolation, LC-ESI-HRMS metabolomic analysis, and LC-Orbitrap-HRMS proteomic analysis. Embryos injected with FMPs can develop into tadpole stage 45 with normal behavior and morphology. Scale bar: 250 µm.

5.3 Experimental section

5.3.1 *Materials and chemicals*

FMPs were from Nvigen (San Jose, CA). Dextran Alexa Fluor 488 fluorescent dye was from Thermo Fisher Scientific (Waltham, MA). Neodymium disk magnets were from Applied Magnets (Plano, TX). Acetylcholine, methionine, threonine, arginine, guanine, sodium citrate, guanidine hydrochloride, and bovine serum albumin (BSA) were purchased from Fisher Scientific (Waltham, MA). HEPES, sodium dodecyl sulfate, sodium isethionate, sodium pyrophosphate, CAPS, proline, and adenosine were from Sigma Aldrich (St. Louis, MO). Citrulline, cysteine, and isoleucine were from MP Biomedicals (Solon, OH). Sodium hydroxide was from Ricca Chemical (Arlington, TX). All solvents including formic acid, water, acetonitrile, methanol, isopropanol, and ethanol were LC-MS grade and from Fisher Scientific (Waltham, MA) unless specified. All chemicals were reagent grade or higher.

5.3.2 *Solutions*

100% Steinberg's solution (58.2 mM NaCl, 0.67 mM KCl, 0.34 mM $Ca(NO_3)_2$, 0.83 mM $MgSO_4$, 4.19 mM Trizma hydrochloride, and 0.66 mM Trizma base, in Milli-Q purified water). Cysteine solution (2% w/v, pH 8). 1 mg/mL BSA solution. Newport dissociation buffer (100 mM sodium isethionate, 20 mM sodium pyrophosphate, and 10 mM CAPS, pH 10.5, in sterile water).[79] 1× Marc’s Modified Ringers (MMR) buffer (100 mM NaCl, 2 mM KCl, 1 mM $MgSO_4$, 2 mM $CaCl_2$, 5 mM HEPES, and 0.1 mM EDTA, pH 7.4, in Milli-Q purified water). 1× phosphate-buffered saline (PBS) buffer. 1× PBS buffer with 2% (w/v) BSA.

5.3.3 *Animal care and X. laevis embryo preparation*

All animal care and handling complied with the requirements and regulations of the Institutional Animal Care and Use Committee (IACUC) of the University of Maryland (IACUC

No. R-FEB-21-07). The behavioral assay was performed under IACUC No. R-JUN-20-31. *X. laevis* frog embryos were obtained through the natural mating of adult frogs induced by gonadotropin.[141] The embryo jelly coatings were removed using 2% cysteine (pH 8). 2-cell stage embryos that displayed stereotypical cleavage and pigmentation were cultured to 16-cell embryos. Embryos showing stereotypic cleavage patterns and pigmentation were used for microinjection.

5.3.4 Needle preparation and microinjection of magnetic particles

Borosilicate glass capillaries (O.D./I.D.: 1.0 mm/0.50 mm, B1005010, Sutter Instrument) were pulled by a micropipette puller (P-1000, Sutter Instrument) using the following settings: heat=495; pull=30; velocity=60; time=150; pressure=200. The pulled needles were prepared under a microscope (SMZ1270, Nikon) by cutting off the tip to have an opening ~8 μm to ensure the filling of the suspension containing green FMPs (2 mg/mL, Cat# 41002, Nvigen). To calibrate the injection volume of needles, needles were filled with sterile water and calibrated to ensure an injection volume of 1 nL per injection using a microinjector (PLI-100A, Warner).[141] Needles were emptied and refilled with 2 μL magnetic particle suspension. 16-cell embryos were placed in an injection plate filled with 100% Steinberg's solution with 3% (w/v) ficoll. Magnetic particles with diameters ranging from 200 to 500 nm were intracellularly microinjected (PLI-100A, Warner) into both left and right D11 (for neural-fated cell clones) or V11 (for epidermal-fated cell clones) blastomeres of 16-cell embryos using a pulled glass needle (P-1000, Sutter Instrument). Each blastomere was microinjected with 4 nL of 2 mg/mL magnetic particle solution.

5.3.5 Behavioral and anatomy assay

Embryos with both left and right D11 cells injected with 4 nL of 2 mg/mL magnetic particles in 1× PBS (Exp group) were cultured to tadpole stage 45 following our established protocol[26] where they have developed visual function. Similarly, both left and right D11 cells injected with 1× PBS were cultured to tadpole stage 45 as the Veh group. Wild-type embryos without injections were cultured to tadpole stage 45 as the Ctrl group. The positive control group (+Ctrl) was obtained by double axotomy of the optic nerve of wild-type stage 45 tadpoles. The behavioral assay was performed to monitor the time percentage that a tadpole spent on the light background side in a tank following our established protocol[26]. The anatomic assay was performed by imaging tadpoles from the Exp and Ctrl groups, followed by measuring the size of tadpoles and the location/size of neural tissues.

5.3.6 Embryo dissociation and cell suspension preparation

Embryos injected with magnetic particles were cultured in a 14°C incubator (Precision, Thermo Fisher Scientific) to five key embryonic development stages: blastula (Nieuwkoop-Faber, NF stage 8), gastrula (NF stage 10 and 12), and neurula (NF stage 13 and 15). To perform embryo dissociation[79], the vitelline membrane of embryos was first removed in 50% Steinberg's solution, followed by washing these embryos with ice-cold 1× MMR solution. Then, embryos were transferred to a BSA-pretreated 12-well plate (Corning 3513, Fisher Scientific) filled with Newport dissociation buffer (5 embryos per well). To minimize cell lysis during dissociation, wells were sealed using parafilm to prevent the formation of air bubbles during vortexing. Plates were vortexed (Genie 2, Scientific Industries) for 5 min on setting 3, for 15 min on setting 5, and for 5 min on setting 7[79], followed by placing them on ice for 10 min to let cells settle down. The dissociation solution was removed as much as possible, followed by adding 200 μL of 1× PBS with 2% BSA to each well immediately. For each sample, cell suspensions from 2 wells (i.e., ~10 embryos) were aspirated into a 1-mL syringe (1001TLL, Hamilton). Each well was rinsed using 100 μL of 1× PBS with 2% BSA to maximize the harvesting of dissociated cells and the wash solution was aspirated into the 1 mL syringe as well. To minimize cell lysis when dissociating neurula stage 8 embryos, the dissociation procedure was revised to vortex for 5 min on setting 1 and 5 min on setting 2. The dissociated samples were placed on ice for 10 min to let cells settle down, followed by removing most of the dissociation solution, adding 400 μL of 1× PBS with 2% BSA per well, and placing on ice for 5 min. Then, most of the solution in each well was removed and 100 μL of 1× PBS with 2% BSA was added. Then, the obtained cell suspension and the wash solution were transferred to 1-mL syringes.

5.3.7 Magnetic cell isolation to collect tissue-specific cells

The 1-mL syringes containing the cell suspension were connected to a borosilicate glass capillary (O.D./I.D.: 1.0 mm/0.75 mm, B1007510, Sutter Instrument) using a plastic sleeve. The glass capillaries were positioned perpendicularly to avoid the accumulation of unlabeled cells in the capillary during magnetic isolation, thus improving the sorting purity. Ten neodymium disk magnets were packed together and placed as close as possible to the glass capillary. Cell suspensions were infused using a syringe pump (Pump 11 Pico Plus Elite, Harvard) at a flow rate of 20 μL/min until the syringe was empty. Cells labeled with magnetic particles were trapped in the glass capillary under the magnetic force. Then, the capillaries were removed gently and checked under an inverted microscope (Eclipse Ti-U, Nikon) to find the collected cell clones under the guidance of green fluorescent signals. The capillary section containing green fluorescence-labeled cells was cut off using a diamond knife, followed by flushing using ~500 μL 1× PBS to wash out the collected cells in 1.5-mL Eppendorf vials. To maximize the recovery rate of labeled cell clones, 100 μL of 1× PBS with 2% BSA was used to rinse the syringe and the solution underwent the same magnetic cell isolation procedure to collect the remaining labeled cells. The collected samples were combined and centrifuged at 400 ×g for 5 min at 4°C. The supernatants were disposed and 100 μL 1× PBS was added to the pellets to wash off BSA before centrifuging again at 400 ×g for 5 min at 4°C. The collected cell pellets were added with 200 μL ice-cold methanol to quench metabolism. Samples were stored in a –80°C freezer immediately until sample preparation to isolate metabolites and proteins for multi-omic analysis. Each biological replicate was from 10 FMP-injected embryos (called MagCar D11) and a total number of N=10 biological replicates were collected.

5.3.8 Metabolomic sample preparation: metabolite extraction

Metabolites were extracted from isolated cell clones using a mixture of 40% (v/v) acetonitrile (ACN) and 40% (v/v) methanol in water. The extraction residues were stored in –80°C freezer until sample preparation for proteomics and transcriptomics sample preparation. Metabolite extract supernatants were reconstituted in 50 μL of 95% (v/v) acetonitrile in water, followed by centrifuging at 13,000 ×g for 10 min at 4°C. The supernatants were transferred into HPLC vials for LC-ESI-MS analysis.

5.3.9 Proteomic sample preparation

Following metabolite extraction, the remaining pellet for each sample was used for protein extraction as follows. Samples were randomized into two batches, which were prepared independently. Pellets were resuspended in 100 μL lysis buffer (1% sodium dodecyl sulfate (SDS), 150 mM NaCl, 5 mM ethylenediaminetetraacetic acid (EDTA), 20 mM Tris-HCl) and aspirated several times to lyse cells and extract proteins. Note: the volumes described in this section were used for the dissected samples; volumes were reduced 10× for FACS and MagCar samples unless otherwise noted as these sorting methods result in less extracted protein. Extracted proteins were then reduced with 4 µL 0.5 M dithiothreitol (DTT) at 60°C for 30 min and alkylated with 12 µL 0.5 M iodoacetamide for 20 min in the dark. The reaction was quenched with the same volume of DTT and was immediately vortexed. Protein was precipitated with chilled acetone at –20°C overnight and pelleted with centrifugation at 10,000 ×g at 4°C for 10 min. After removing supernatants, protein pellets were reconstituted in 50 mM ammonium bicarbonate (adjusted to pH 8 with Tris) and trypsin was added (4 µg for dissected samples, 0.75 µg for FACS/MagCar samples) and proteins were digested at 37°C for 12 h. The resulting peptides were dried by centrifugal evaporation. Peptide amounts in each sample were quantified using a Take3 micro-volume plate (Agilent Technologies, Santa Clara, CA). A pooled sample was created from equimolar amounts of all samples for quality control (QC) purposes. One pooled sample was labeled as part of each batch of samples with nine tags of Tandem Mass Tags (TMT)^{10}-plex (Thermo Fisher Scientific, Waltham, MA; 25 µg per sample). Each sample was reconstituted to 1 µg/µL in 100 mM triethylammonium bicarbonate and incubated with 0.2 mg of a randomly assigned TMT reagent for 1 h with vortexing. The reaction was quenched by adding 2 µL 5% hydroxylamine to each sample and incubating for 15 min with vortexing. The labeled samples from each batch were combined and each batch mixture was dried by centrifugal evaporation. Each batch mixture (100 µg) was separated into 72 fractions using high pH reversed-phase fractionation on an Agilent 1260 Infinity II liquid chromatography (LC) system equipped with a fraction collector (Agilent). Labeled peptide mixtures were loaded onto an InfinitiLab Poroshell HPH-C18 guard column (Agilent; 4.6×5 mm, 2.7 µm) and then separated on an InfinitiLab Poroshell HPH-C18 column (Agilent; 4.6×150 mm, 2.7 µm) at 0.5 mL/min over the following gradient: 0–17 min, 0% B; 17–18 min, 0–7% B; 18–75 min, 7–35% B; 75–76 min, 35–100% B; 76–80 min, 100% B; 80–86 min, 100–0% B; 86–90 min, 0% B. Mobile phases A and B were 10 mM ammonium bicarbonate in water and 90% ACN, respectively, both adjusted to pH 10 with ammonium hydroxide. Fractions were only collected between 17–85 min. Fractions were dried via centrifugal evaporation and concatenated into 8 final fractions. Each fraction was reconstituted in 0.1% formic acid (FA) for LC-tandem mass spectrometry with MS^3 (MS^3) analysis.

5.3.10 Metabolomic analysis using LC-ESI-MS

The samples were queued in an autosampler, which was programmed to inject 2 µL samples into a BEH Amide Column (Waters; 1 mm×100 mm, 1.7 µm). Metabolites were separated by the Waters ACQUITY I-Class UPLC system at 130 µL/min and 40°C column temperature over the following 10 min gradient: 0–0.5 min, 95% B; 0.5–10 min, 95–40% B; 10–13 min, 40% B; 13–15 min, 40–95% B; 15–22 min, 95% B. Mobile phases A and B were 0.1% FA in LC-MS grade water and ACN, respectively. The separated metabolites were ionized by an ESI source, and detected by a time-of-flight mass spectrometer (timsTOF Pro, Bruker). We measured samples in the mass range *m/z* 20–1,000 under positive mode. Automated data-dependent acquisition (DDA) collision-induced dissociation (CID) MS/MS spectra were acquired to identify metabolites from isolated cell clones.

5.3.11 Proteomic analysis using LC-MS^3

Fractions were injected in a random order in technical triplicate (1 µg) for LC-MS^3 analysis on an UltiMate™ 3000 RSLCnano system (Thermo Fisher Scientific) coupled to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) operated in positive mode. Peptides were loaded onto an Acclaim PepMap™ 100 C18 trap column (Thermo, 100 µm×2 cm, 100 Å, 5 µm) at 5 µL/min and then were separated on a µPAC™ column (PharmaFluidics, Ghent, Belgium; 200 cm) at 0.600 μL/min and 50°C over the following 240 min gradient: 0–5 min, 1% B; 5–20 min, 1–7% B; 20–135 min, 7–25% B; 135–160 min, 25–32% B; 160–193 min, 32–45% B; 193–200 min, 45–75% B; 200–208 min, 75% B; 208–210 min, 75–2% B; 210–240 min, 2% B. Mobile phases A and B were 0.1% FA in LC-MS grade water and ACN, respectively. Full MS spectra were collected in the Orbitrap (400–1600 *m/z*, 120,000 resolution). The mass spectrometer was operated in data-dependent acquisition (DDA) mode with a 2.5 s cycle time to collect MS/MS scans in the ion trap using collision-induced dissociation (CID; 35% normalized collision energy, 0.7 *m/z* isolation window, turbo ion trap scan rate) with a dynamic exclusion of 60 s. The peptide monoisotopic peak determination node was on and only precursor ions with charges +2–+6 and an intensity >5.0×10^3 were considered for MS/MS fragmentation. MS^3 spectra were then collected using synchronous precursor selection (SPS) to simultaneously fragment the top 10 most intense fragments from the MS/MS spectra in the Orbitrap using higher-energy C-trap dissociation (HCD; 100–500 *m/z*, 55% normalized collision energy, 50,000 resolution, 0.7 *m/z* MS isolation window, 3 *m/z* MS/MS isolation window). Precursor ions were only selected from the 400–1600 *m/z* range for MS^3 with a ±25 ppm exclusion mass width; precursor ions with a TMT tag loss were excluded from selection for MS^3. All spectra were collected using the standard automatic gain control (AGC) target and auto maximum injection time mode.

5.3.12 Metabolomic data processing

The obtained MS metabolomic data files were processed using Bruker MetaboScape Version 4.0.4 software to find molecular features (unique *m/z* and retention time) and corresponding intensity. Metabolite identification was based on accurate mass, retention time, and MS/MS matching[182] against the METLIN database[11] and MassBank of North America (MoNA) database[183] allowing for highly confident identification. Peak intensity was used as a quantitative proxy to compare relative metabolite level changes. Student's *t*-tests were conducted to identify metabolites with statistical significance ($p<0.05$). MetaboAnalyst 5.0[224] was used to conduct metabolomic data normalization, statistical analyses, and pathway analyses.

5.3.13 Proteomic data processing

Proteome Discoverer software (version 2.2) was used to analyze RAW files. All fractions and technical replicates for each batch were combined into one file for analysis. Each batch file was searched against a concatenated database of *X. laevis* proteins including proteins from Xenbase[139] and the mRNA-derived Proteomic Reference with Heterogeneous RNA Omitting the Genome (PHROG[162]), as well as a database of common contaminants, using SEQUEST-HT. Modifications included in the search were carbamidomethylation on cysteine residues as a fixed modification and TMT^{10}-plex (229.163 Da) on peptide N-termini and lysine residues, oxidation of methionine residues, and acetylation of the protein N-termini as dynamic modifications. Two trypsin missed cleavages were allowed in the search. Mass tolerances were 10 ppm for precursor ions and 0.6 Da for fragment ions. Peptides were required to have a false discovery rate (FDR) of ≤1%.

TMT^{10}-plex quantification methods were used for both batches, each of which was customized to only include the 9 tags used for that batch. TMT reporter ion abundances were based on intensity with a signal-to-noise ratio of ≥30; quantification results were corrected based on the isotopic impurities of the TMT lot used in these experiments. All quantification results were normalized to the total peptide amount in Proteome Discoverer. Protein identifications in each batch were filtered to only include Master Proteins with FDR ≤1% and ≥2 peptide spectral matches (PSMs). Contaminant proteins were excluded from further analysis. Quantified proteins were required to be identified in both batches and have TMT reporter ion intensities for both pooled samples and ≥3 biological replicates per group. Most proteins filtered out at this step were missing reporter ion intensities in the FACS and MagCar groups, while fewer were missing values from the Dissect groups. This is likely due to the presence of BSA in the FACS and MagCar samples from the sorting buffers, which can interfere with the detection and quantification of low-abundance proteins. The intensities for quantified proteins were median-normalized and $\log_{10}$ transformed using MetaboAnalyst 5.0[224]. Normalized data were used to calculate fold changes and *p*-values using two-tailed Student's *t*-tests between sample groups. Fold-change cutoffs were calculated for each group comparison based on technical and biological variation and replication (Diss FMP vs Diss WT, |fold change|≥1.34; FACS vs Diss, |fold change|≥1.47; MagCar vs Diss FMP, |fold change|≥1.66)[184]. Differentially expressed proteins had $p < 0.05$ and met the defined fold-change cutoffs. Interactions among differentially expressed proteins were predicted using STRING version 11.5[185]. Gene Ontology annotation and statistical overrepresentation analyses were performed for differentially expressed proteins against *Xenopus tropicalis* using the PANTHER database[186] version 17.0 (released 02/22/2022).

5.3.14 Image scoring to study the effect of magnetic particles on cell fate determination of injected cells

Injection needles were prepared as previously described. To explore the effect of injected magnetic particles on neural tissue development, the right D11 cells on 16-cell embryos were injected with 4 nL of a mixture of 2 mg/mL magnetic particles and 50 pg/nL *gfp* mRNA as the experimental group while the right D11 cells on 16-cell embryos were injected with 4 nL of 50 pg/nL *gfp* mRNA only as the control group. To explore the effect of injected magnetic particles on epidermal tissue development, the right V11 cells on 16-cell embryos were injected with 4 nL of a mixture of 2 mg/mL magnetic particles and 50 pg/nL *gfp* mRNA as the experimental group while the right V11 cells on 16-cell embryos were injected with 4 nL of 50 pg/nL *gfp* mRNA only as the control group. The injected embryos were cultured under room temperature (21°C) to larva stage 34, followed by fixation using 4% paraformaldehyde with 3% sucrose in 1× PBS. The fixed larvae were imaged using a stereomicroscope (SMZ18, Nikon) and the obtained images were scored for neural and epidermal tissues following a previously established protocol[6].

5.4 Results and discussion

5.4.1 MagCar development and mathematic validation

Isolating tissue-specific cells from developing embryos is crucial for studying cell differentiation and organogenesis. Generally, tissue-specific cell clones (descendant cells originated from the same precursor cells) from developing organisms have rapidly changing cell sizes, a large number of descendant cells, and varying locations due to cell migration. These factors all make targeting and isolating tissue-specific cell clones for downstream analysis more complex. Therefore, to address this bottleneck, we developed a workflow called MagCar to incorporate intracellular microinjection of FMPs followed by cell lineage tracing and embryonic dissociation, with magnetic force-based isolation to collect tissue-specific cell clones (**Fig. 5.1**). This strategy enabled efficient isolation of targeted cell clones with a wide range of cell sizes from developing vertebrate organisms (*Xenopus* frog) and could be potentially applicable to various models with established cell fate maps including zebrafish. Here, we used *X. laevis* embryos as an example to demonstrate the application of MagCar to understand tissue formation by isolating neural- and epidermal-specific cell clones from five key embryonic stages to study temporal metabolomic and proteomic changes.

5.4.2 Cell lineage tracing enabled by MagCar

We microinjected FMPs into precursor cells to track and label tissue-specific cell clones over embryonic development using the setup in **Fig. 5.2A**. To minimize the toxicity of FMPs, we titrated the microinjected quantity of FMPs while ensuring that injected magnetic particles can label and track targeted cell clones during embryonic development. To demonstrate the cell lineage tracing capability of MagCar, we used D11 cell clones as an example. We injected green FMPs into both left and right D11 cells of 16-cell embryos (**Fig. 5.2B**), followed by confocal microscopy imaging of the injected FMPs in D11 cells (**Fig. 5.2C**). The image results showed that FMPs inside D11 cells barely aggregated, indicating that FMPs were dispersed to efficiently label D11 cell clones at later stages during embryonic development. The green fluorescence property of the injected FMPs enabled cell lineage tracing by fluorescence imaging under a fluorescence microscope. **Fig. 5.2D** showed D11 cell clones at the gastrula stage labeled by FMPs as highlighted in the white line. To demonstrate that FMPs labeled not only the surface D11 cell clones but also D11 cell clones underneath the embryo's surface[22], we further bisected embryos labeled by FMPs at the gastrula stage and fluorescence microscopy imaging showed that D11 cell clones underneath the embryo surface were labeled by magnetic particles (**Fig. 5.2E**). Thus, the intracellular microinjection of FMPs was demonstrated to enable the cell lineage tracing of targeted cell clones over development. To evaluate the potential aggregation of magnetic particles after being injected into cells, we used a confocal microscope to image cells labeled by FMPs, which provided information about the sizes of magnetic particles inside cells. We plotted histograms for magnetic particle size in solution, in D11 cells, and in D11 cell clones at the gastrula stage. The histogram plots showed that magnetic particle size had a similar distribution pattern before and after being injected into cells, which revealed that injected magnetic particles did not aggregate during cell lineage tracing. Therefore, MagCar enabled efficient labeling and tracking of targeted cell clones during development.

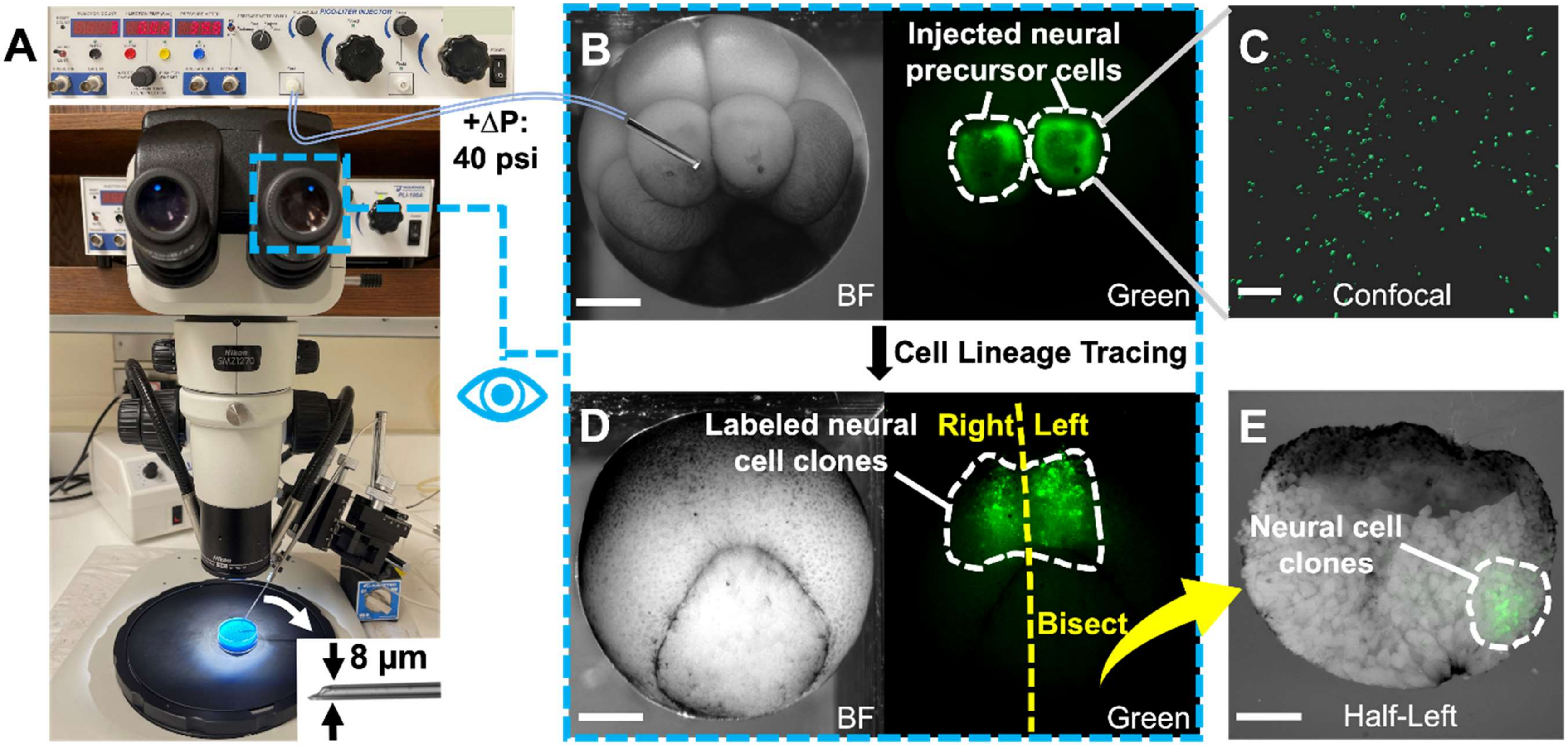


Figure 5.2 Cell clone labeling by MagCar. **(A)** Setup for the intracellular microinjection of fluorescent magnetic particles. The setup consists of a microinjector, microscope, three-axis motorized stage, a capillary holder, and a clay-coated dish. Images were embryos injected with fluorescent magnetic particles in the right D11 cell. Inserted image exemplified the capillaries used for FMP injection. **(B)** bright-field (left) and fluorescent (right) images of 16-cell stage embryos injected with green FMPs. Scale bar: 250 µm. **(C)** Confocal image of D11 cell. Green dots were FMPs inside D11 cells. Scale bar: 15 µm. **(D)** Bright-field (left) and fluorescent (right) images of a gastrula-stage embryo injected with magnetic particles. The highlighted area is D11 cell clones labeled by FMPs. Scale bar: 250 µm. **(E)** Image of half-left bisected embryo from **(D)**. Scale bar: 250 µm.

5.4.3 Normal embryonic development with injected magnetic particles: morphological and behavioral assay

To use MagCar to isolate tissue-specific cell clones, a prerequisite was to ensure the injected magnetic particles did not adversely affect the stereotypical development of embryos. We demonstrated that the injected magnetic particles caused no delay in embryonic development (**Fig. 5.3A**). By tracking the development timeline, we found embryos injected with magnetic particles developed synchronically from cleavage stage 5 to tadpole stage 45 compared to wild-type embryos. For tadpoles that reached stage 45, we conducted morphological assays of tadpoles' size (**Fig. 5.3B**). When comparing the size of wild-type tadpoles (N=5) with tadpoles from FMP-injected embryos (N=5), the Mann-Whitney test returned $p=0.531$, $p=0.060$, $p=0.095$, and $p=0.403$ for body length, center-to-center eye distance, center-to-center olfactory distance, and eye location, respectively (**Fig. 5.3C**). A further morphological assay of neural tissue was performed because the injected D11 cells were neural-fated precursor cells (**Fig. 5.3D**). When comparing the area of neural tissues between wild-type tadpoles and tadpoles from FMP-injected embryos, the Mann-Whitney test returned $p=0.095$, $p=0.676$, and $p=0.835$ for eyes, olfactory, and brain, respectively (**Fig. 5.3E**). These statistical analyses revealed that the injected magnetic particles did not cause statistically significant ($p<0.05$) morphological change to injected embryos during embryonic development.

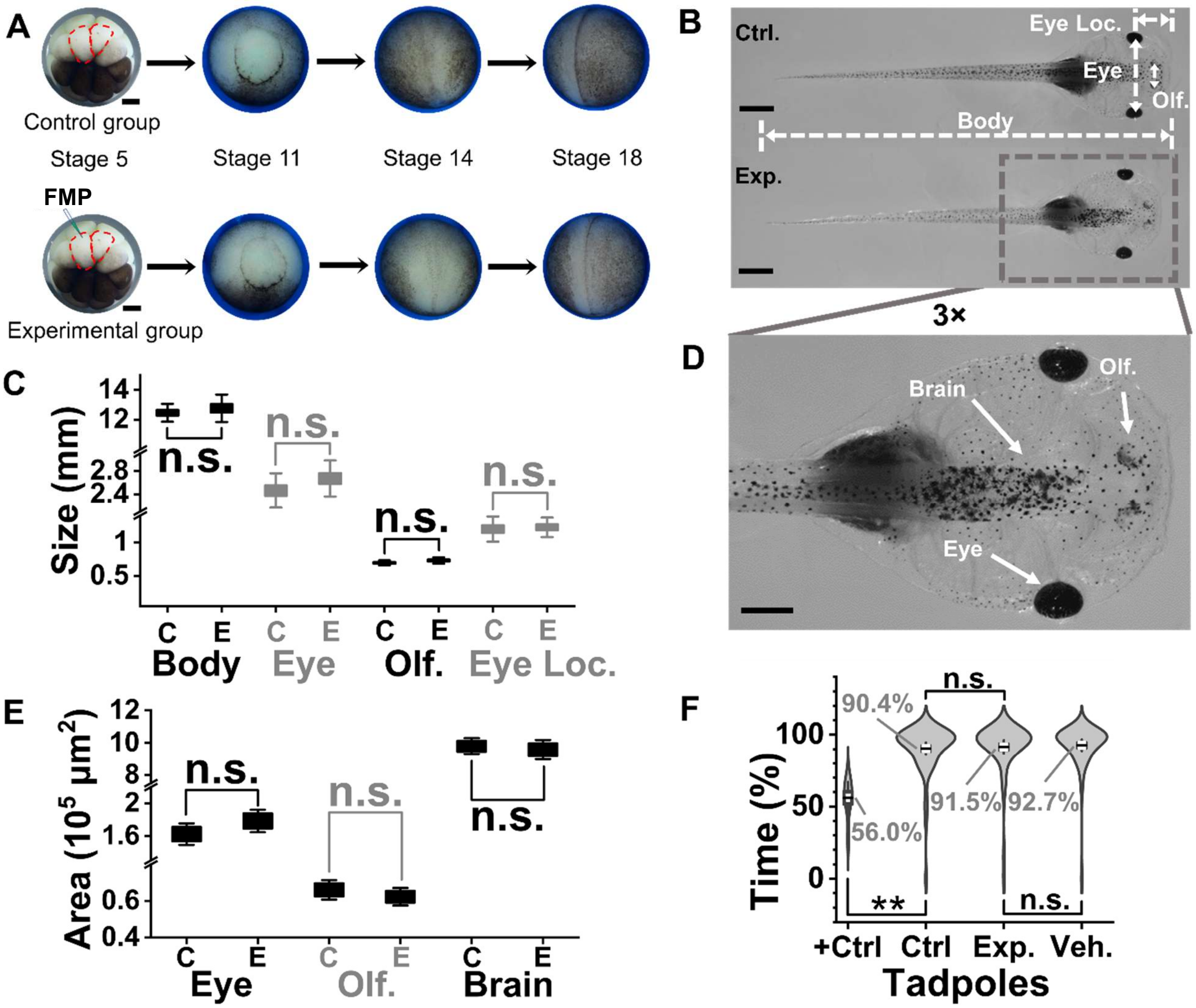


Figure 5.3 Evaluation of the effect of injected magnetic particles on embryonic development. **(A)** Top: images of wild-type tadpoles; Bottom: images of embryos injected with fluorescent magnetic particles (FMP) at different developmental stages. Each group contains N=10 embryos. Scale bar: 250 μm. **(B)** Anatomical and morphological analysis of tadpole sizes between wildtype tadpoles (C, N=5) and tadpoles from embryos with both left and right 2D11 injected with FMPs (E, N=5) tadpoles. Body length, center-to-center eye distance, center-to-center olfactory distance, and eye location were selected for analysis. Scale bar: 1 mm. **(C)** Mann-Whitney test returned p=0.531, p=0.060, p=0.095, and p=0.403 for body length, center-to-center eye distance, center-to-center olfactory distance, and eye location, respectively. **(D)** Anatomical and morphological analysis of neural tissues between wildtype tadpoles (C, N=5) and tadpoles from embryos with both left and right D11 injected with FMPs (E, N=5) tadpoles. Area of eyes, olfactory, and brain were selected for analysis. Scale bar: 0.5 mm. **(E)** Mann-Whitney test returned p=0.095, p=0.676, and p=0.835 for eyes, olfactory, and brain, respectively. **(F)** Behavioral assay (background color preference assay) of stage 45 tadpoles. +Ctrl (tadpoles with double axotomy of the optic nerves, N=4); Ctrl (wild-type tadpoles, N=15); Exp. (tadpoles from embryos with left and right D11 cells injected with magnetic particles, N=14); Veh. (tadpoles from embryos with left and right D11 cells injected with 1×PBS injected, N=13). Key: n.s., non-significant; *p<0.05, **p<0.005 (Mann-Whitney).

However, similar morphology did not necessarily mean similar behavior. To investigate if magnetic particles cause any behavioral changes, we conducted a behavioral assay called background color preference assay for *X. laevis* stage 45 tadpoles labeled with/without magnetic particles by adopting the previous protocols.[26, 143] This assay evaluated the visual function, signal processing, and moving capability of tadpoles. Since we microinjected FMPs into D11 cells which were the progenitors of the neural tissues including the retina, brain, and spinal cord, we would expect to observe behavioral changes if the injected magnetic particles interfered with the visual and neural functions of tadpoles from FMP-injected embryos. We first validated the behavioral assay by comparing the Ctrl group (N=15, wild-type tadpoles) with the +Ctrl group (N=4, tadpoles with double axotomy of the optic nerves) and found the +Ctrl group spent half of the time on the light background side (56%), indicating the +Ctrl group tadpoles had no visual function to sense background color difference and therefore swam randomly. In contrast, the wild-type tadpoles spent 90.4% of their time on the white background side, and the wild-type tadpoles' preference for the white background matched with previous findings[26, 143]. The Mann-Whitney test between the Ctrl and the +Ctrl group revealed a significant difference (p=3.046×10^{-5}) between these two groups (**Fig. 5.3F**). This meant the behavioral assay that we conducted was sensitive to detect behavior change caused by impaired visual function. When comparing the Exp. group (N=14, tadpoles from FMP-injected embryos) with the Ctrl group (N=15, wild-type tadpoles), statistical results revealed that the Exp. group preferred the light background side, spending 91.5% of the time on the white side like their wild-type siblings and was indistinguishable from the Ctrl. Group (p=0.498, Mann-Whitney test) (**Fig. 5.3F**). These results confirmed that magnetic particle injection did not interfere with normal embryonic development, specifically causing no developmental delay, morphological change, or behavioral change, which was supported by findings from the previous studies[225-226].

5.4.4 Mathematical validation of the feasibility of MagCar cell isolation setup

To apply MagCar to isolate cells labeled by magnetic particles from developing embryos, we designed a magnetic cell sorting apparatus (**Fig. 5.4A**). This setup consists of a programmable pump, a syringe, a glass capillary, and disk magnets. During the magnetic cell isolation, cell suspensions were supplied at a steady flow rate of 20 µL/min by a syringe connected to a pump. Cells containing magnetic particles were trapped in the capillary under magnets whereas unlabeled cells were flushed out of the capillary. To avoid the clogging of the sorted cells in the glass capillaries, we chose commercially available and cost-effective glass capillaries with a large internal diameter of 750 µm. Capillaries with a large diameter allow the isolation of cells with diameters up to ~200 µm which was far beyond the capability of the current standard FACS instruments, making MagCar attractive for sorting cells with relatively large sizes. The glass capillaries were positioned vertically during sorting so that unlabeled cells were flushed out of the capillaries by dragging force and gravity, helping to improve sorting purity.

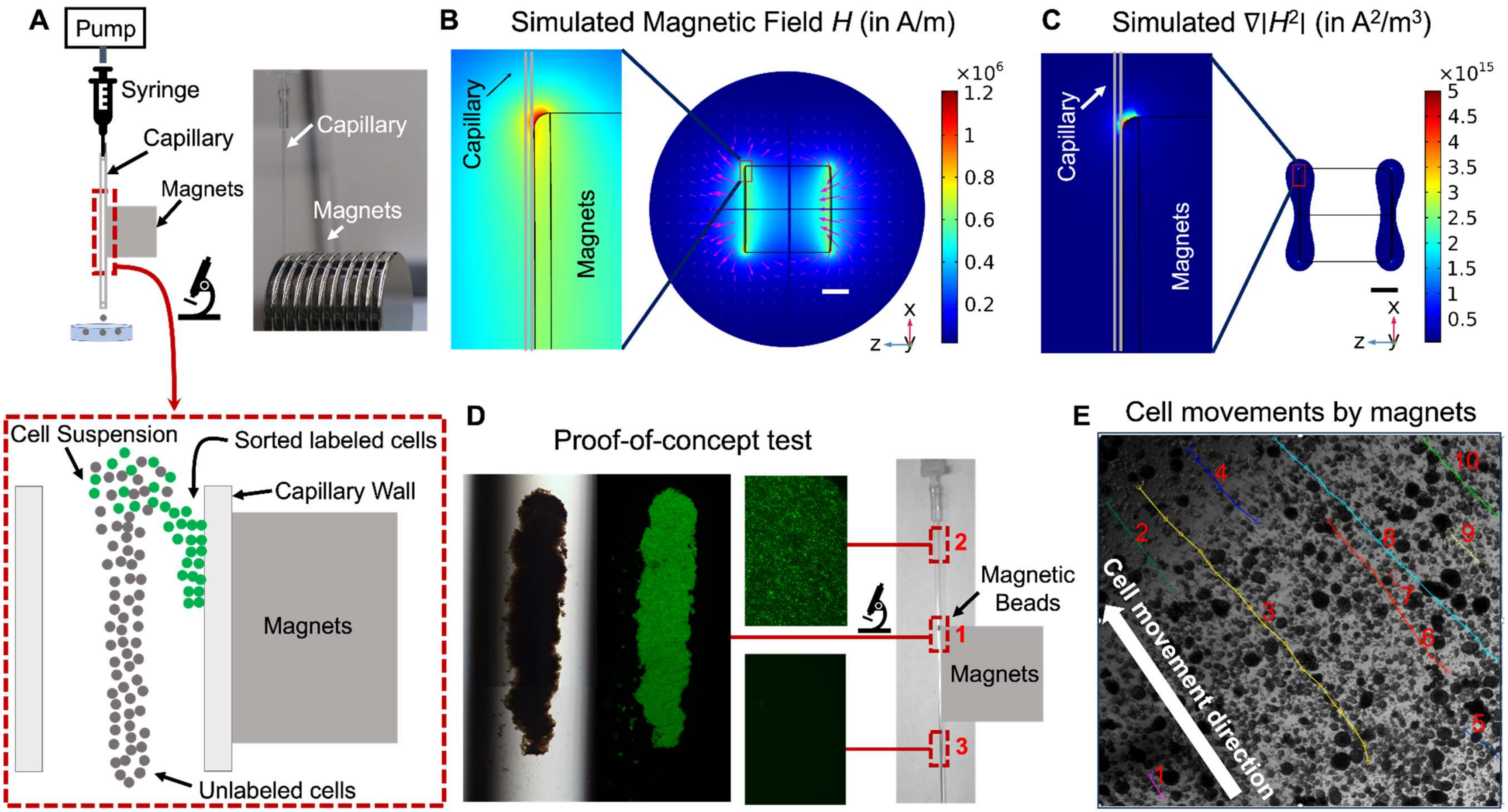


Figure 5.4 Validation of the feasibility of MagCar to isolate tissue-specific cell clones. **(A)** Scheme of designed sorting setup which consists of a syringe collected to a programmable pump, a syringe, a glass capillary, and magnets. Green dots stand for cells labeled by green FMPs and these cells are collected during sorting; grey dots stand for unlabeled cells and these cells are discarded during sorting. **(B)** Simulated magnetic field *H* of neodymium magnets. The arrows show the direction of the magnetic field. Scale bar: 10 mm. **(C)** Simulated magnitude of $|H^2|$ gradient around magnets. A threshold of 3×10^{13} A^2/m^3 was selected to plot $\nabla|H^2|$. Scale bar: 10 mm. **(D)** Proof-of-concept test. 1: the fluorescent image of magnetic particles trapped in the capillary during magnetic isolation; 2: the fluorescent image of magnetic particles suspension before magnetic isolation; 3: the fluorescent image of droplets eluted from the glass capillary during magnetic isolation. **(E)** Cell movement trajectories under the magnetic force. Color lines were the selected 10 trajectories of cell movement obtained using ImageJ software.

To ensure that our setup provided enough magnetic force to isolate cells with magnetic particles, we mathematically calculated the magnetic force required to overcome the drag force experienced by a cell, thus, cells containing magnetic particles can be isolated from unlabeled cells. According to the Stokes equation: $F_d = 6\pi\eta a v$ (where η is viscosity of fluid, a is the diameter of a cell, υ is the velocity of cell movement), the drag force experienced by a single cell at the gastrula stage (a diameter of 50 μm) was 356 pN under a flow rate of 20 μL/min. To calculate the magnetic force experienced by a single magnetic particle in our sorting apparatus, we simulated the magnetic field of the disk magnets in our setup using the COMSOL Multiphysics 5.5 software given the relative permittivity (μ_r=1.05) and remnant field strength (B_r=1.45 T) for the neodymium magnets that we used. The simulation result demonstrated that the magnetic field strength throughout the glass capillary was greater than 0.5 T (**Fig. 5.4B**). This simulated magnetic field result was experimentally confirmed using a magnetic field meter (MF100, Extech) which read a value of 5,229 Gauss (i.e., 0.5229 T) at the location where the glass capillary was placed. A magnetic field strength of > 0.5 T means all superparamagnetic particles were saturated[227] when cells containing magnetic particles passed through the capillaries. Therefore, we met the requirement to use **Equation 1** to calculate the magnetic force F_m that a single superparamagnetic particle experienced inside the glass capillary:

$$F_m = 2\pi r^3 \mu_0 \mu_r K \nabla |H^2| \qquad \textbf{Equation 1}$$

Where r is the radius of a magnetic particle, μ_0 is the permeability of free space, μr is the relative permeability of fluid, H is the magnetic field, $\mu_{r,p}$ is the relative permeability of the magnetic particle. K is defined in **Equation 2** as:

$$K = \frac{\mu_{r,p} - \mu_r}{\mu_{r,p} + 2\mu_r} \qquad \textbf{Equation 2}$$

Simulation results in Figure 3C showed that $\nabla|H^2|$ was greater than 3×10^{13} A^2/m^3 throughout the entire capillary. Assuming $\nabla|H^2|=3\times10^{13}$ A^2/m^3, a single magnetic particle (radius r=100 nm) experienced a magnetic force of 0.237 pN according to **Equation 1**. This value was comparable to the magnetic force calculated from the previous studies[228-229]. However, at the edge of the magnets, the $\nabla|H^2|$ can be as high as 2.5×10^{15} A^2/m^3 (**Fig. 5.4C**), indicating a magnetic force of 19.7 pN experienced by a single magnetic particle according to **Equation 1**. Therefore, to ensure that cells contain enormous magnetic particles to generate large enough magnetic force, we chose to microinject the precursor cells on 16-cell embryos with 4 nL of magnetic particles (a concentration of 1×10^8 particles/μL, N=4×10^5). Each of the descendant cells inherited ~6,000 particles at the gastrula stage assuming cells cleaved equally at each stage. Therefore, the total magnetic force experienced by a single cell at the gastrula stage could be as high as 123 nN, which was 2–3 orders of magnitude compared to the drag force. Therefore, the magnetic cell isolation setup of MagCar provided sufficient magnetic force to overcome the drag force so that cell clones labeled by magnetic particles could be isolated with minimal sample loss.

To experimentally validate the proposed magnetic isolation setup, we performed a proof-of-concept test using magnetic particle suspension. When the magnetic particle suspension was pushed through the glass capillary with a flow rate of 20 μL/min, we observed that magnetic particles were trapped exactly at the edge of the interface between magnets and the glass capillary as shown in **Fig. 5.4D**. This matched with the simulation results showing the greatest value of $\nabla|H^2|$—the strongest magnetic force according to **Equation 1** (**Fig. 5.4C**). Furthermore, no magnetic particles were found on the capillary eluent during magnetic isolation, which supported that our setup isolated magnetic particles efficiently.

For cells labeled by magnetic particles, we visualized cell movement under the magnetic force using a microscope. Through this experiment, we demonstrated that the position of cells containing FMPs could be manipulated using disk magnets. Further investigation of cell movement trajectories revealed that all tracked trajectories (shown in **Fig. 5.4E**) were parallel and pointed to the location of magnets, which demonstrated that cell movements indeed resulted from magnetic force and that the magnetic force was strong enough to manipulate cell movements. Thus, magnetic cell isolation combined with magnetic particle injection and embryo dissociation was feasible to collect tissue-specific cells from developing embryos.

5.4.5 Evaluate the performance of MagCar magnetic cell isolation apparatus

To experimentally test the magnetic cell isolation apparatus of MagCar, we prepared a cell suspension mixture by dissociating neurula-stage embryos cultured from 16-cell stage embryos with both left and right D11 cells injected with FMP. The obtained cell suspension contained both FMP-labeled cells and unlabeled cells (**Fig. 5.5A**). The cell circled in red color was labeled by FMP. According to the trypan blue exclusion test, >95% of cells in the dissociated cell mixture were viable. Using our magnetic isolation setup, we visually observed that cells were trapped in the glass capillary by magnets as highlighted in the red rectangle in **Fig. 5.5B**. We imaged the highlighted region under an inverted microscope (Eclipse Ti-U, Nikon) and found that cells retained in the capillary were fluorescently labeled (**Fig. 5.5C**). To accurately determine the sorting purity, we imaged the isolated cells under the 20× confocal microscope (LSM 980 Airyscan 2, Zeiss) with higher sensitivity compared to that of the inverted microscope. For more than 500 MagCar isolated cells, we found that 92.4% of them were labeled by clear green FMP signals (**Fig. 5.5D**). However, the actual purity could be even higher because some cells might be labeled but the fluorescence signals were below the lower limit of detection of optical microscopy. Compared with the traditional MACS techniques having a purity range from 64–95%[216], our MagCar workflow's isolation purity of 92.4% was on the higher side. A closer look at the MagCar-isolated cells under the 63× confocal microscope revealed that the green FMPs were distributed randomly inside cells with no preferred subcellular locations (**Fig. 5.5E**). Moreover, the cell membrane of MagCar sorted cells were intact, which meant that the MagCar workflow provided sufficient magnetic force to isolate cells labeled with FMPs without damaging the integrity of sorted cells during the sorting process. The preservation of cell viability during MagCar was further confirmed by the trypan blue exclusion test which showed that 96.3% (1,282/1,331) of MagCar collected cells were viable. Further investigation of the fluorescence profile of these collected cells revealed a normal distribution pattern of the fluorescence intensity (**Fig. 5.5F**). To evaluate the potential sample loss during MagCar, we collected the eluent during the magnetic sorting process and imaged the collected solution. Imaging analysis showed that < 0.3% of cells with green fluorescent signals were eluted. Thus, the MagCar sorting apparatus enabled the isolation of the FMP-labeled cells with high purity, high viability, and minimal sample loss.

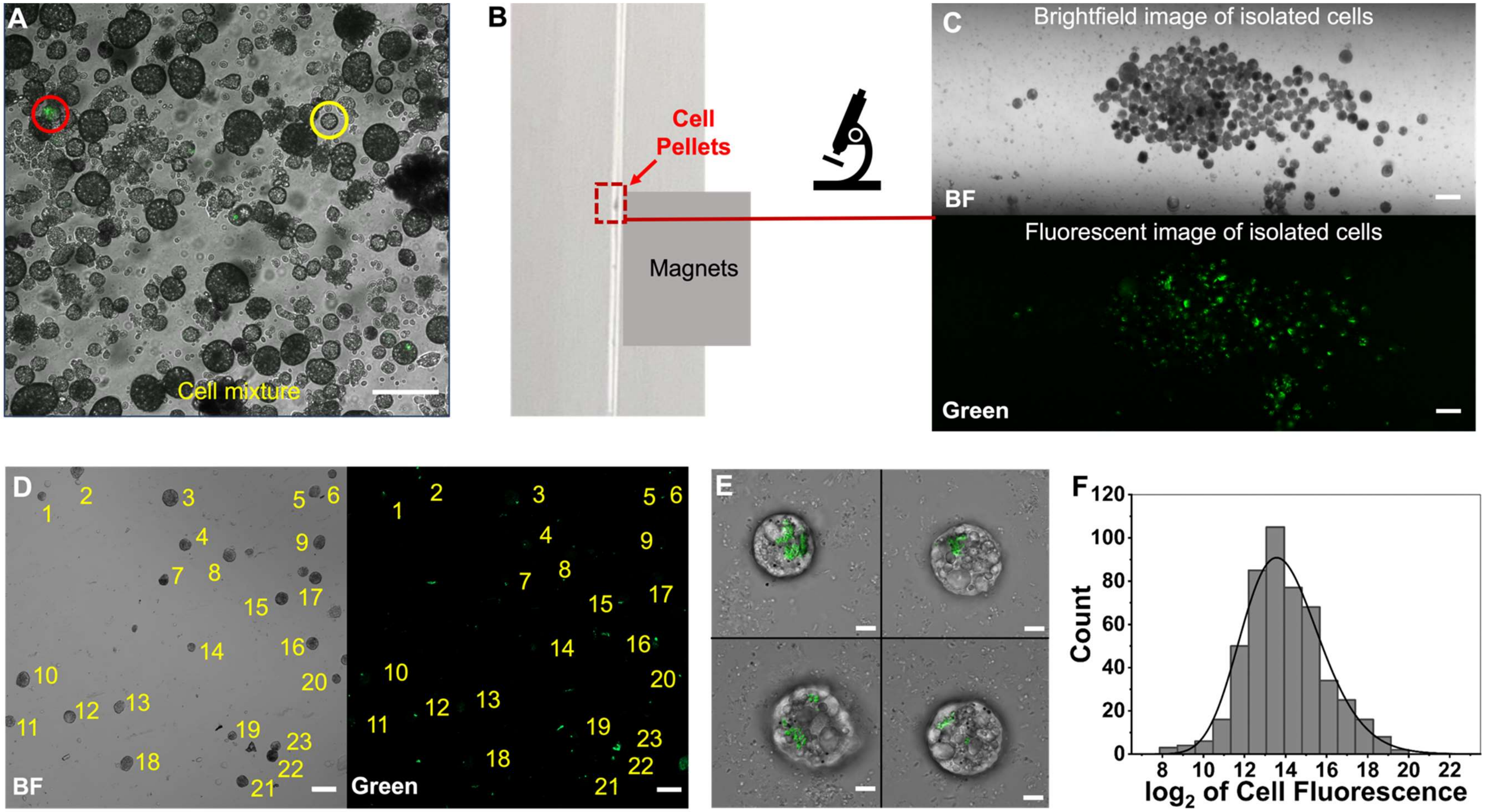


Figure 5.5 MagCar to collect tissue-specific cell clones using a custom-built sorting apparatus. **(A)** Image of cell mixture from dissociated neurula stage 13 embryos injected with magnetic particles. Cell labeled by green fluorescent magnetic particles was exemplified in a red circle. Unlabeled cell was exemplified in a yellow circle. Scale bar: 50 µm. **(B)** D11 cell clones isolated by designed apparatus. The collected cell clones were highlighted in a red rectangle**.** **(C)** Images of collected cell clones under an inverted microscope. Top: brightfield (BF) image of collected cell clones; Bottom: green fluorescence (Green) image of collected cell clones. Scale bar: 50 µm. **(D)** Brightfield (left) and green fluorescence (right) images of MagCar sorted D11 cell clones using confocal microscope under 20× magnificence. The digits note cells with magnetic particles. Scale bar: 50 µm. **(E)** Overlapped brightfield and green fluorescence images of MagCar sorted D11 cell clones using confocal microscope under 63× magnificence. The magnetic particles were visible and distributed randomly. Scale bar: 5 µm. **(F)** Histogram to show the fluorescence intensity distribution of sorted ~500 D11 cell clones. The inserted curve showed the normal distribution.

5.4.6 Study the molecular change of labeled cell clones caused by injected magnetic particles

To investigate whether injected magnetic particles cause molecular perturbations, we used dissected neural stage 13 D11 cell clones labeled by dextran green (called Diss WT group, N=10 biological replicates, **Fig. 5.6A**) as a reference of native states. This is because dextran dye is biologically inert and resistant to endogenous enzymes.[141] We injected a mixture of dextran green and magnetic particles into D11 cells and dissected out the labeled D11 cell clones at neurula stage 13 (called Diss FMP group, N=10 biological replicates, **Fig. 5.6A**). Metabolomic and proteomic analyses were both performed for the two dissected groups (**Fig. 5.6B**).

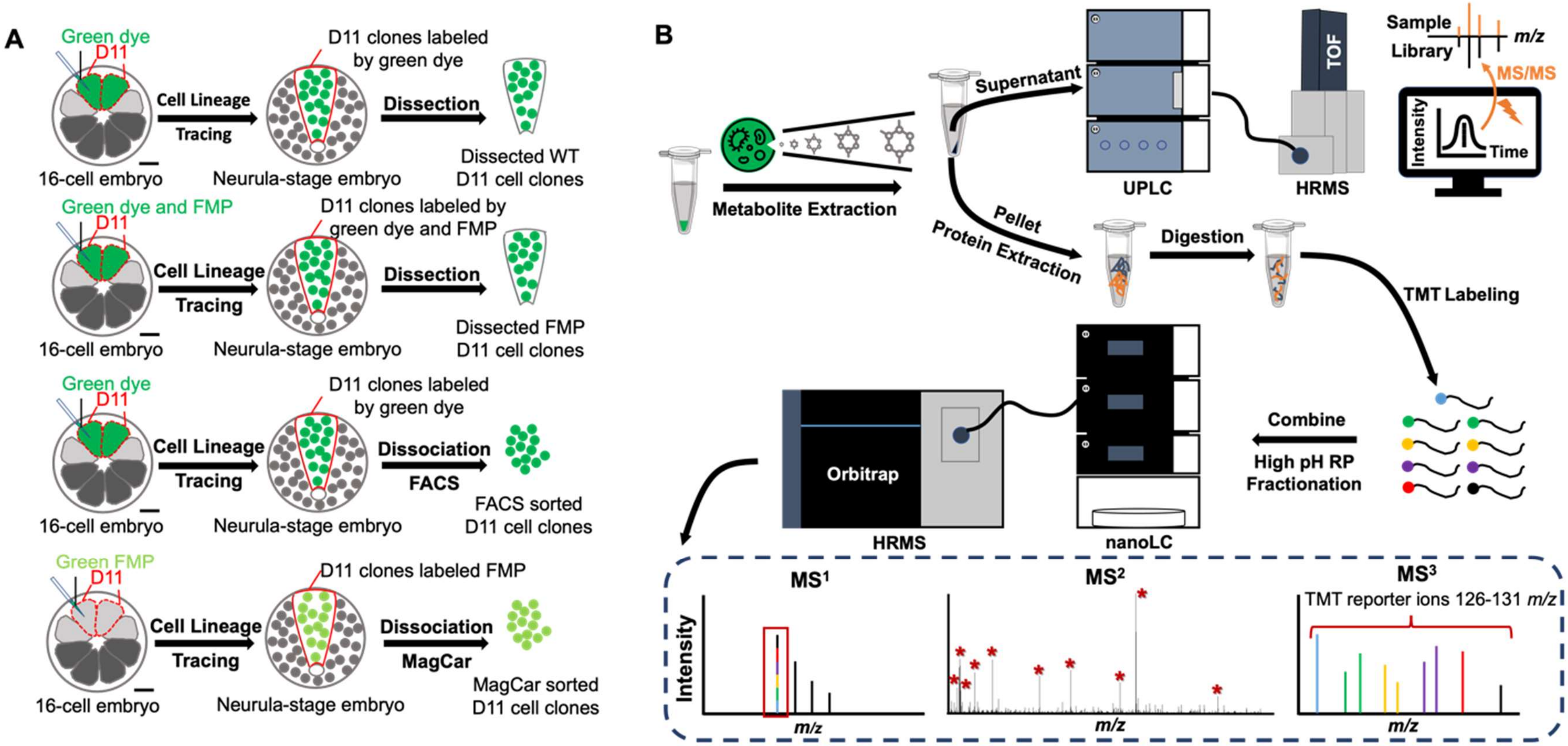


Figure 5.6 Sample collection and preparation to evaluate the metabolic and proteomic perturbations during MagCar using LC-HRMS. (**A**) Workflow to dissect wild-type neurula stage 13 D11 cell clones labeled by green fluorescent dye (called Diss WT, N=10); workflow to dissect neurula stage 13 D11 cell clones labeled by green magnetic particles and fluorescent dye (called Diss FMP, N=10); workflow to isolate neurula stage 13 D11 cell clones by FACS (called FACS, N=6); workflow to isolate neurula stage 13 D11 cell clones by MagCar (called MagCar, N=10). (**B**) Workflow of dual metabolomic and proteomic analysis of the collected samples.

We conducted the power analysis of the detected molecular features in untargeted metabolomic analysis (**Fig. 5.7A**). The predicted power was 5.46×10^{-3} between the Diss WT group and the Diss FMP group (N=10) for the metabolomic data, which was far below 0.8 (a typical threshold to confidently distinguish two groups). Even if the sample number was increased to 1,000, the predicted power was still below 0.8 and could not differentiate the Diss WT group from the Diss FMP group with statistical confidence. Since GSH/GSSG ratio is widely used as a marker to evaluate the oxidative stress experienced by cells,[230] we adopted here to use GSH/GSSG ratio to study if the injected FMP caused extra oxidative stress to D11 cell clones. The Mann-Whitney test of GSH/GSSG ratios between the Diss WT group (N=10) and the Diss FMP group (N=10) returned a *p*-value of 0.186 (**Fig. 5.7B**) indicating that D11 cell clones experienced no statistically significant difference in oxidative stress after being injected with FMPs.

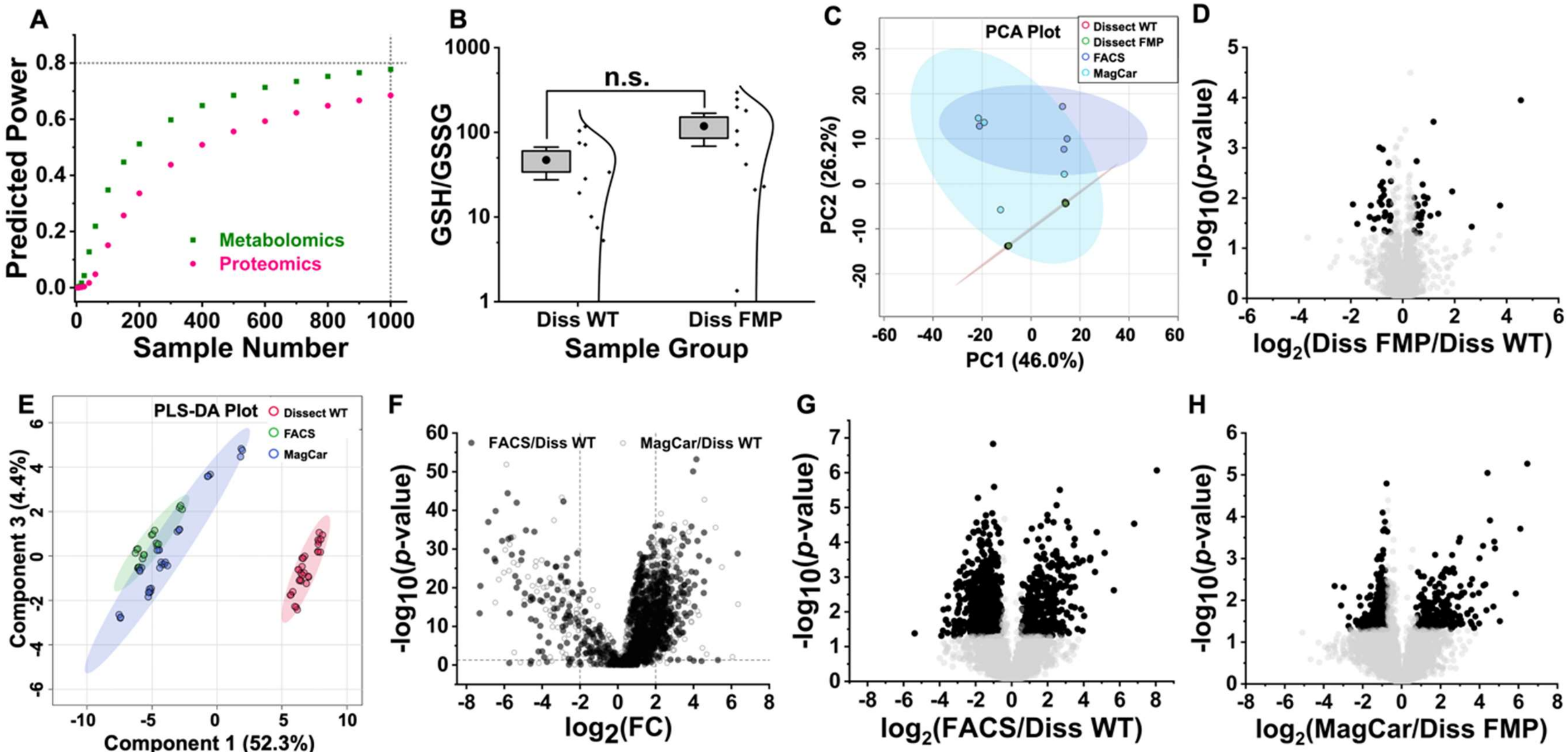


Figure 5.7 Evaluation results of potential metabolic and proteomic perturbations during MagCar. (**A**) Power analysis of metabolomic and proteomic data between Diss WT and Diss FMP, respectively. Green dots meant expected power at the corresponding sample number using all molecular features removed of contaminant molecular features from magnetic particles; red dots meant expected power at the corresponding sample number using all quantified proteins. When the sample number N=1,000, Diss FMP was chemically indistinguishable from Diss WT (Power <0.8) according to both metabolomic and proteomic data. (**B**) Mann-Whitney test of oxidative stress indicator—GSH/GGSG ratio between Diss WT (N=10) and Diss FMP (N=10). (**C**) Principal component analysis (PCA) of proteomic data for Diss WT (N=4), Diss FMP (N=4), FACS (N=4), and MagCar (N=4). (**D**) Volcano plot between Diss WT and Diss FMP to explore perturbed proteins by injected magnetic particles. Black circles were proteins that showed significant changes ($p<0.05$, $\log_2|\text{fold change}|\geq 0.43$); grey circles were proteins without significant changes. (**E**) Partial least squares-discriminant analysis (PLS-DA) for Diss WT (N=10), FACS (N=6), and MagCar (N=10). (**F**) Volcano plot of metabolomic data between FACS (N=6) and Diss WT (N=10) (dark circles); volcano plot between MagCar (N=10) and Diss WT (N=10). (**G**) Volcano plot of proteomic data between FACS (N=4) and Diss WT (N=4). Black circles were proteins that showed significant changes ($p<0.05$, $\log_2|\text{fold change}|\geq 0.55$); grey circles were proteins without significant changes. (**H**) Volcano plot of proteomic data between MagCar (N=4) and Diss FMP (N=4). Black circles were proteins that showed significant changes ($p<0.05$, $\log_2|\text{fold change}|\geq 0.73$); grey circles were proteins without significant changes.

We further evaluated the potential proteomic change caused by injected FMP by comparing the Diss WT group (N=4) with the Diss FMP group (N=4). Samples were analyzed using a TMT^{10}-plex quantitative proteomics workflow (**Fig. 5.6B**). Across the two TMT batches, a total of 6,112 unique proteins were identified, 3,864 of which were identified in both batches. This subset was further filtered based on the presence of TMT reporter ion intensities (see Methods) to result in 3,517 quantified proteins. When performing power analysis using these quantified 3,517 proteins to compare the Diss WT group with the Diss FMP group, the predicted power was also below 0.8 when the sample number was increased to 1,000 (**Fig. 5.7A**). The low expected power from both 'omic analyses suggested that the Diss WT group was indistinguishable from the Diss FMP group. A principal component analysis (PCA) was also performed (**Fig. 5.7C**). The PCA plot showed overlap across all groups, thus indicating that none of the sample collection methods resulted in distinct proteomic profiles compared to the Diss WT group. The Diss WT and Diss FMP groups overlapped almost exactly, highlighting the similarity in proteomic profiles in these two groups and that injecting the FMPs did not result in significant perturbation to the proteome of these cells. This observation was supported after statistical comparisons between the Diss WT and Diss FMP groups as well, between which only 60 proteins (1.7% of the quantified proteins) were differentially expressed (**Fig. 5.7D**). The majority of these perturbed proteins function in various metabolic pathways, primarily the metabolism of carbohydrates and small molecules. In summary, the above metabolomic and proteomic data leads to the conclusion that the injected FMPs caused only minimal molecular changes in the labeled cell clones.

5.4.7 MagCar mitigates molecular changes experienced during FACS sorting

To assess molecular perturbations in different sample collection methods (**Fig. 5.6A**), we further compared the potential metabolomic and proteomic changes caused by the developed MagCar with that of FACS using the Diss WT D11 cell clones group or Diss FMP D11 cell clones group as a control. We used molecular features detected from metabolite extracts to study the metabolic state of the cells collected by MagCar and compare potential changes to those caused by FACS. According to a supervised statistical analysis called partial least square discriminant analysis (PLS-DA) (**Fig. 5.7E**), the MagCar group (N=10) clustered closer to the Diss WT group (N=10) compared with the FACS group (N=6), indicating that MagCar minimizes potential metabolic changes to collected samples. Furthermore, volcano plots revealed 705 molecular features detected from FACS-sorted neural cell clones change significantly (*p*-value <0.05, fold change≥2) (**Fig. 5.7F**, black circles). Specifically, 113 molecular features are down-regulated, and 592 molecular features are up-regulated (*p*-value <0.05, fold change≥2). The revealed metabolic change of sorted cells during FACS matched with a study reported previously.[219-220] In contrast, fewer molecular features from neural cell clones collected by MagCar undergo significant change. Among these 595 molecular features, 101 molecular features are down-regulated, and 494 molecular features are up-regulated (**Fig. 5.7F**, grey circles). Therefore, the MagCar technique alleviated the metabolic change that happened during the sample collection process.

From the perspective of proteomics, both FACS and MagCar cell isolation caused changes to the proteome of collected cells. However, the FACS group only slightly overlapped with the Diss WT and Diss FMP groups in the PCA plot (**Fig. 5.7C**), indicating that the proteomic profile of the FACS-collected samples was largely different from the dissected samples. Based on the volcano plots (**Fig. 5.7G** and **Fig. 5.7H**), FACS resulted in a more significant impact on the proteome compared to MagCar. Specifically, FACS resulted in almost 30% of the total quantified proteins (N=1,035) having significant differences compared to Diss WT (**Fig. 5.7G**). In contrast, the proteomic impact of MagCar sample collection was reduced to nearly half that of FACS, resulting in 469 perturbed proteins (13.3% of quantified proteins; **Fig. 5.7H**). PANTHER statistical overrepresentation analyses of the proteins perturbed in FACS revealed that translation, protein/peptide biosynthesis, and metabolism were enriched among proteins decreased in FACS samples (**Fig. 5.8A**), while metabolism and protein transport/localization were the primary biological processes characterized by proteins increased in FACS samples (**Fig. 5.8B**). Similar pathways were observed among the proteins perturbed by MagCar (**Fig. 5.9A-B**). However, the overrepresented pathways in proteins increased in the MagCar group primarily only belonged to metabolic pathways (**Figure 5.9B**). In contrast to the FACS samples, proteins in protein transport and localization pathways were not increased in MagCar samples, indicating that FACS resulted in more widespread disruption of proteomic pathways than MagCar.

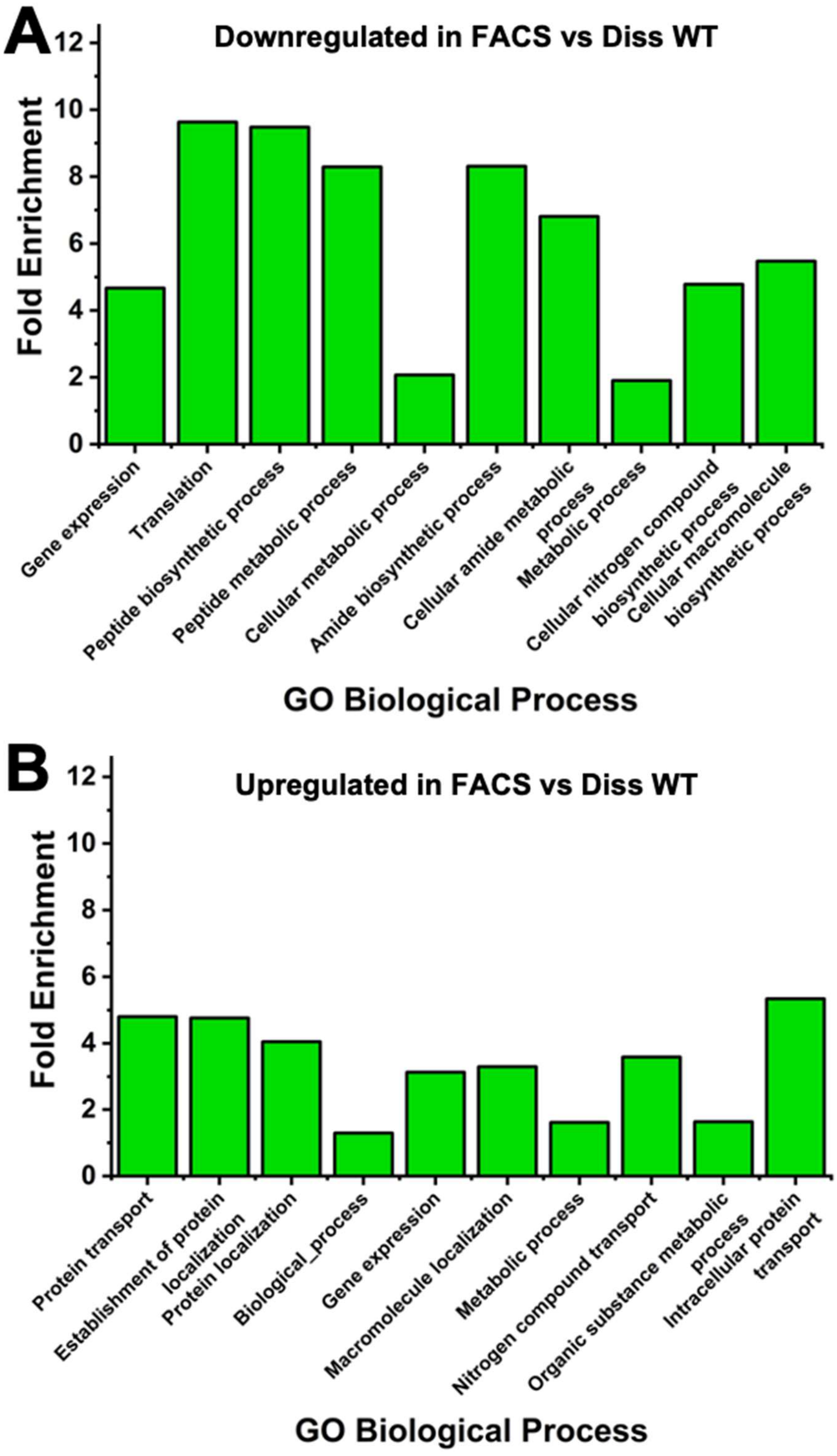


Figure 5.8 Overrepresented biological processes among differentially-expressed proteins in FACS vs Diss WT. Top 10 overrepresented Gene Ontology Biological Processes in **(A)** downregulated and **(B)** upregulated differentially-expressed proteins in FACS vs Diss WT groups. The top 10 biological processes are the 10 with the lowest false discovery rate (FDR) for statistically significant overrepresentation. Overrepresentation, represented by fold enrichment, is relative to the *Xenopus tropicalis* genome. Abbreviations: GO, Gene Ontology; FMP, fluorescent magnetic particles.

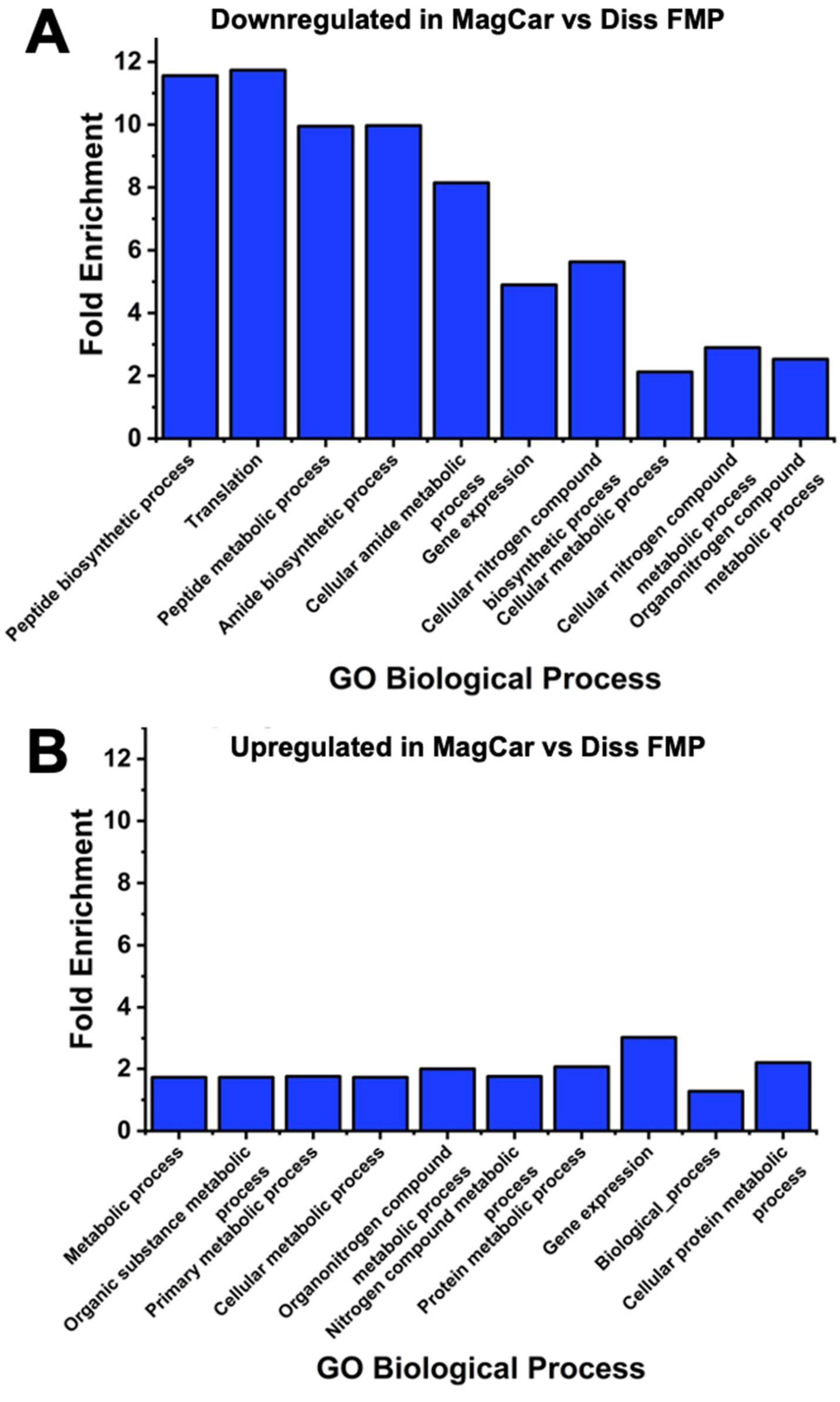


Figure 5.9 Overrepresented biological processes among differentially-expressed proteins in MagCar vs Diss FMP. Top 10 overrepresented Gene Ontology Biological Processes in **(A)** downregulated and **(B)** upregulated differentially-expressed proteins in MagCar vs Diss FMP groups. The top 10 biological processes are the 10 with the lowest false discovery rate (FDR) for statistically significant overrepresentation. Overrepresentation, represented by fold enrichment, is relative to the *Xenopus tropicalis* genome. GO, Gene Ontology; FMP, fluorescent magnetic particles.

Therefore, our developed MagCar technique helps partially mitigate both metabolic and proteomic changes during sample collection compared with FACS. The improvement may be attributed to that the fact that MagCar required no sample transfer between labs, a shorter sample collection time, and was free of electric potential/strong laser light during sorting. We expect that our MagCar technique can further mitigate molecular changes during the sample collection process by using chemically inert or bioinert magnetic particles as well as using a temperature-controlling system during magnetic cell isolation.

5.4.8 MagCar vs FACS for tissue-specific cell isolation

MagCar sorting workflow has advantages in sample collection throughput, scalability, sample recovery rate, and expense, compared with FACS. For *Xenopus* cell suspension (10 to 35 µm in cell diameter) at the neurula stage, MagCar takes ~30 mins to collect 1 sample regardless of cell size. The sample collection throughput of MagCar scales up easily by using parallel setups, allowing for simultaneous cell sorting. In contrast, FACS cannot allow for the simultaneous isolation of multiple samples and requires ~45–60 mins per sample. This does not count the time to filter dissociated cells before FACS and transport time to/from the sorting facility. The sorting time of FACS can increase substantially for cells with large diameters due to system clogging. For a developing system with continuous cell division such as *Xenopus* in which cells cleave every ~45 min[139], faster sampling ensures the isolation of targeted cell clones before the next cleavage cycle begins.

MagCar has better scalability to isolate cells ranging from a few micrometers to hundreds of micrometers, whereas typical FACS instruments are limited to sorting smaller cells due to clogging of the tubing when sorting large cells (>40 µm). This is because the workable cell size of a typical FACS instrument is limited to ¼ of the nozzle size, typically 130 µm in diameter, to avoid clogging issues. As a comparison, the sorting capillary in MagCar has a diameter of 750 μm which enables to sort cells with diameters ~150 μm as demonstrated for cells from *Xenopus* blastula stage 8. The capability to sort these large cells expands the application of MagCar to isolate tissue-specific cell clones from early-stage embryos/organisms.

MagCar allows the isolation of targeted cell clones with a higher sample recovery rate. Our total peptide assay revealed a higher amount of peptide in MagCar-isolated cell clones compared with that in FACS-isolated cell clones. Furthermore, the metabolomic data shows a higher abundance (~3 times) of metabolite signals in MagCar-collected cell clones than that of FACS under the same starting material condition. The lower sample recovery rate in FACS may result from sample loss during sample preparation in which dissociated cell suspensions undergo a filtration step before FACS experiments. Another possibility is that cells need to pass through a narrow nozzle during FACS experiments which may cause cell lysis. Moreover, cells could lyse during transportation to and from the sorting facility. All these factors could contribute to the lower sample recovery rate from FACS.

Furthermore, MagCar enables in-house sample collection without using an expensive and sophisticated instrument and avoids the need for instrument maintenance. The MagCar workflow only involves syringes, a syringe pump or pressurized cell suspension supporting system, magnets, and glass capillaries, which means MagCar can be conducted in a regular lab without having to be in a sorting core facility. High school students can perform sorting experiments without needing any in-depth and specialized training. FACS instruments, on the other hand, are expensive and complex and require expertise and in-depth training to conduct sorting, troubleshooting, and maintenance.

In summary, compared with FACS, MagCar is a sampling workflow with advantages in preserving samples' native state, sample collection throughput, scalability, expense, and sample recovery rate while maintaining comparable sorting purity, cell viability, and sample loss.

5.4.9 Explore to apply MagCar to study dorsal-ventral specification

To apply MagCar to understand how a single progenitor cell develops into pre-destined cell clones over embryonic development, the prerequisite is to test whether the intracellularly injected FMPs in the early embryonic cells cause cell fate change. To evaluate the effect of FMPs on the phenotype (cell fate) of injected progenitor cells—neural-fated dorsal progenitor D11 cells, we designed a workflow that combines embryo injection, cell lineage tracing, and image scoring. For the control group, we injected the right D11 cells in 16-cell stage embryos with *gfp* mRNA only to label the progenies of D11 cells. For the experimental group, we injected the right D11 cells in 16-cell stage embryos with a mixture of *gfp* mRNA and FMPs. The injected embryos were cultured to larva stage 34, fixed using 4% paraformaldehyde, and scored following the guidelines in previous publications.[6] Mann-Whitney tests were performed to evaluate the tissue scoring difference between the control and the experimental group. Our results showed that there was no significant difference between the groups by scoring the neural tissue and epidermal tissues. This demonstrated the injected FMPs did not interfere with the normal cell differentiation of neural-fated progenitor D11 cells (**Fig. 5.10A**). Similarly, we applied the strategy to epidermal-fated V11 cells and validated that the injected magnetic particles did not interfere with the normal cell differentiation of epidermal-fated progenitor V11 cells (**Fig. 5.10B**). These findings paved the road to applying the developed MagCar strategy to study tissue formation during embryonic development.

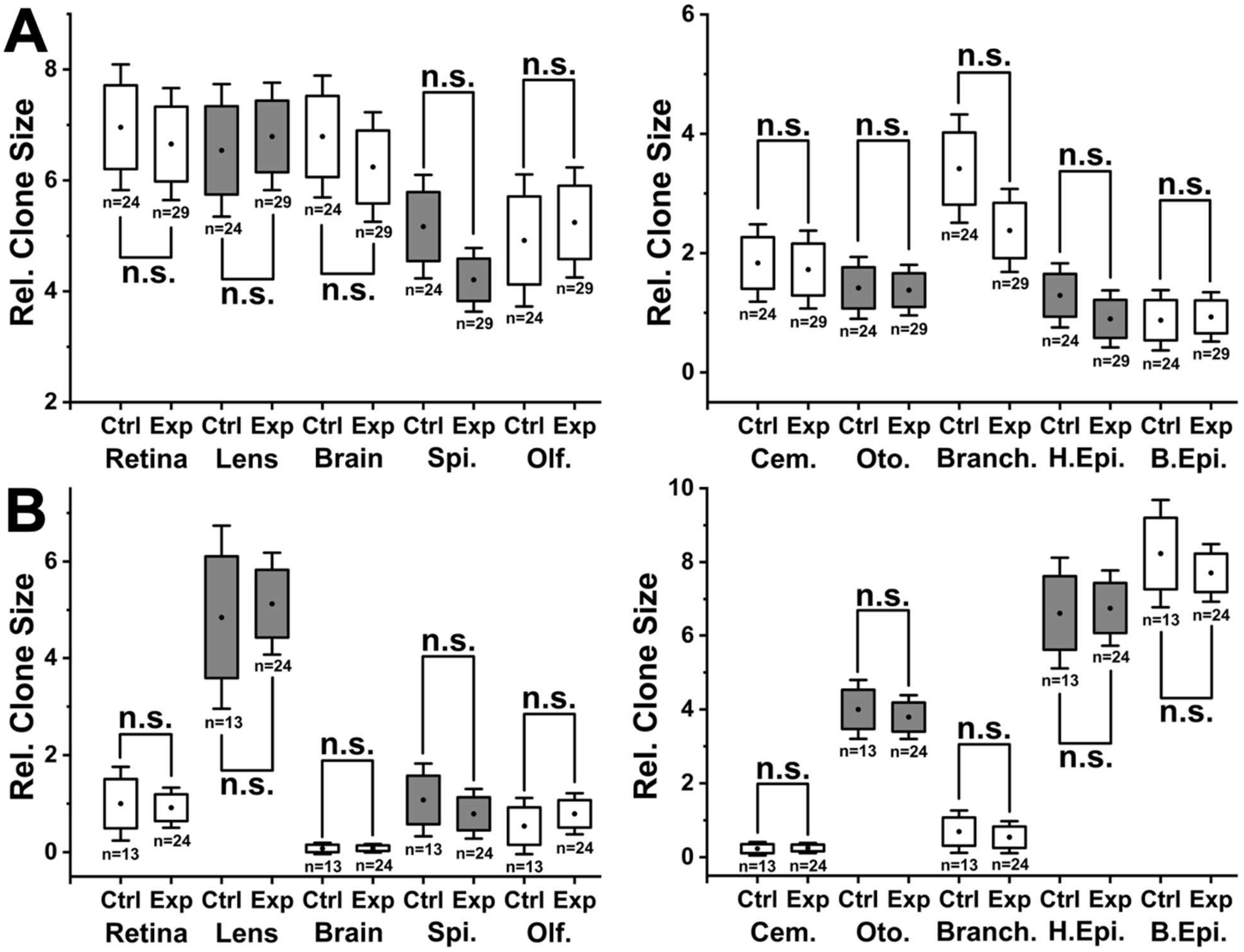


Figure 5.10 Study the effect of injected FMP on cell differentiation and tissue formation. (**A**) Image scoring to study the effect of injected FMP on the phenotype of D11 (R)—precursor cell of neural tissues. Ctrl (N=24): stage 34 larvae developed from 16-cell embryos with D11 (R) injected with *gfp* mRNA only. Exp (N=29): stage 34 larvae developed from 16-cell embryos with D11 (R) injected with *gfp* mRNA and FMPs. *p*-value of tissues based on Mann-Whitney test: Retina (*p*=0.930); Lens (*p*=0.529); brain (*p*=0.725); C.S. (*p*=0.414); Olf. (*p*=0.420); Cem. (*p*=0.446); Oto. (*p*=0.303); Branch. (*p*=0.570); H.Epi. (*p*=0.203); B.Epi. (*p*=0.445). (**B**) Image scoring to study the effect of injected FMP on the phenotype of V11 (R)—precursor cell of epidermal tissues. Ctrl (N=13): stage 34 larvae developed from 16-cell embryos with V11 (R) injected with *gfp* mRNA only. Exp (N=24): stage 34 larvae developed from 16-cell embryos with V11 (R) injected with *gfp* mRNA and FMPs. *p*-value of tissues based on Mann-Whitney test: Retina (*p*=0.063); Lens (*p*=0.220); brain (*p*=0.502); C.S. (*p*=0.745); Olf. (*p*=0.557); Cem. (*p*=0.950); Oto. (*p*=0.719); Branch. (*p*=0.536); H.Epi. (*p*=0.240); B.Epi. (*p*=0.431).

The developed MagCar technique was demonstrated to successfully isolate neural and epidermal cell clones from crucial developmental stages including the blastula stage, gastrula stage, and neurula stage with varied cell sizes (10 to 150 µm in diameter). **Fig. 5.11** shows the isolated neural cell clones using MagCar over blastula NF stage 8, gastrula NF stage 10 and 12,

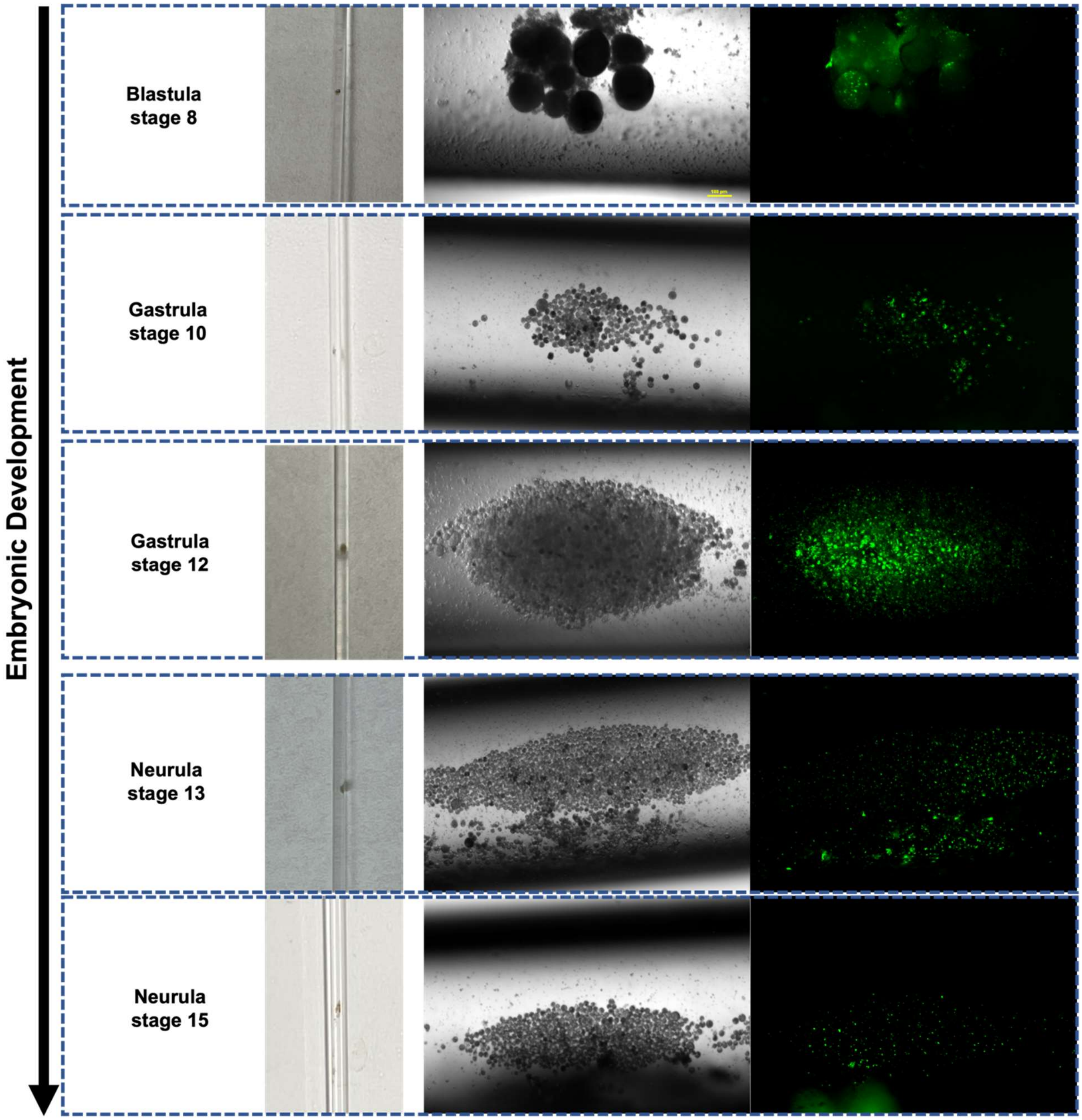


Figure 5.11 Examples of MagCar collected neural cell clones from different embryonic development stages including blastula NF 8, gastrula NF 10, gastrula NF 12, neurula NF 13, and neurula NF 15. Cells at blastula stage 8, gastrula stage 10, and gastrula stage 12 have cell sizes >50 µm, which are too large to be sorted using typical FACS instruments.

and neurula NF stage 13 and 15. Moreover, cells before neurula stage 13 having a diameter of >50 µm are too large to be sorted using a typical FACS instrument. Therefore, the isolation of tissue-specific cells from these early stages is inaccessible previously. However, the developed

MagCar technique addresses this technique gap, opening a new door to studying tissue formation from early to late embryonic stages. Specifically, we are applying MagCar to isolate tissue-specific cells to explore the temporal metabolic and proteomic changes over a wide range of key embryonic development stages along with LC-MS-based multi-omics, which promises to deepen our understanding of tissue formation.

5.5 Conclusions

We have developed a novel sampling technique called MagCar that enables the efficient isolation of tissue-specific cells from developing embryos to address unmet sampling challenges in cell and developmental biology. We have validated that the injected magnetic particles did not interfere with normal embryonic development by morphological, functional, and behavioral assay. Specifically, there was no developmental delay, morphology change, cell fate change, and behavior change. Compared with the FACS, the MagCar technique has advantages in sample collection throughput, scalability, sample recovery rate, and expense. For example, MagCar enabled the isolation of cells with diameters up to ~150 μm which is far beyond the capability of the current standard FACS instruments. Moreover, MagCar maintains comparable sorting metrics to FACS in terms of sorting purity, cell viability, and sample loss. MagCar was demonstrated to isolate high-purity (92.4%) cell clones with high viability (96.3%) and negligible sample loss. We further confirmed that the developed MagCar workflow minimized potential metabolic and proteomic changes to collected samples, compared with FACS. Furthermore, MagCar is readily adaptable to other technologies and workflows, which was demonstrated in this dissertation that MagCar combined with LC-HRMS enabled the metabolomic and proteomic analysis of tissue-specific cell clones to study cell clones' phenotypic states. The MagCar technique we developed in this project can be adopted for other important biological models such as zebrafish to elucidate how cells develop into tissues, organs, and whole organisms during health and disease.

Chapter 6: Conclusions and future directions

6.1 Deepening the understanding of tissue formation

My dissertation research goal was to study an important topic in cell and developmental biology: understanding the biomolecular components of single embryonic cells and exploring their functional roles in cell differentiation to tissues. The insights gained from this research and the developed techniques could potentially be used to understand healthy/dysfunctional tissue development and elucidate the cause of congenital diseases. To uncover the molecular mechanisms underlying tissue formation, it is crucial to understand the molecular differences among single cells with different tissue fates at early embryonic stages and reveal how these molecular differences contribute to the respective tissues during embryonic development. Due to technological limitations, however, only one type of biomolecule has typically been measured in previous studies, which limits our understanding of cell-to-cell differences. There is a significant unmet need to develop technologies for the dual analysis of proteins and metabolites at the single-cell level. Furthermore, it was challenging to use existing techniques to track the development of a cell and specifically isolate its descendant cells at later stages for downstream metabolomic and proteomic analysis. These technological limitations precluded us from gaining a deeper understanding of cell differentiation and tissue formation.

In this dissertation, I developed techniques for cell and developmental biology. I summarized the techniques that we developed for advancing functional and systems biology research in MS-based single-cell analysis (Chapter 2). In Chapter 3, I developed a technique called *in vivo* microsampling, which enabled multiple sampling of the same embryonic cells in live developing embryos without interfering with normal embryonic development. To explore the role of metabolites in cell tissue fate determination, I developed a sampling workflow that combined microinjection, cell lineage tracing, embryonic dissociation, and FACS, which addressed the challenges in isolating targeted tissue-specific cells from early vertebrate embryos with relatively small cell size for multi-omic analysis (Chapter 4). This FACS-based sampling workflow enabled the isolation of tissue-specific cells from neurula-stage embryos. The samples collected from the developed sampling technique were compatible with downstream LC-MS analysis, which helped to reveal molecular remodeling underlying metabolite-induced cell fate change. Moreover, to conduct temporal studies of tissue formation to further deepen the understanding of this biological process, I developed a novel sampling technique called MagCar, which overcomes the challenges in targeting and isolating tissue-specific cells from different embryonic stages with cell sizes ranging from several microns to hundreds of microns (Chapter 5).

In this work, I developed and applied *in vivo* microsampling to analyze the left dorsal-animal (L-D1) cells and left ventral-animal cells (L-V1) in 8-cell embryos. I demonstrated the developed *in vivo* microsampling did not interfere with normal embryonic development. Using the developed *in vivo* microsampling technique and CE-MS, I achieved, for the first time, multi-omic (proteomic and metabolomic) analysis of the same single cell in live embryos developing into normally behaving tadpoles. I revealed molecular cell heterogeneity among early-stage pluripotent cells with different tissue fates along the dorsal-ventral axis by using our multi-omic strategy, which showcased the potential of this strategy for obtaining new biological knowledge. The developed techniques and knowledge gained opened a new door to investigating various systems and developmental biology questions about embryonic development. For example, the *in vivo* microsampling technique developed from this dissertation work can be adopted to investigate subcellular heterogeneity and explore the relationship between molecular abundance and cell fate decision.

Moreover, I studied methionine-induced cell fate change by integrating the developed FACS sampling technique with LC-MS-based metabolomics and proteomics. I revealed perturbed metabolic pathways using metabolomics and proteomics datasets, thus, uncovering molecular remodeling underlying methionine-induced cell fate actions. Specifically, I established the molecular interaction network resulting from the injected methionine, which revealed potential molecular effectors that were involved in cell fate change actions. In the follow-up heavy isotope labeling experiments, I validated the revealed molecular effectors and further proved these molecular effectors caused D11 cells to change cell fate from neural to epidermal tissues. These results improved our understanding of metabolite-induced cell fate change and allowed us to generate hypothesized mechanisms, which will guide the direction of future research to elucidate the role of metabolites in driving cell differentiation. For example, we can conduct metabolic flux analysis to track the consumption of methionine, which will provide crucial information for a better understanding of methionine-induced cell fate change.

My dissertation work also led to the development of a novel sampling technique called MagCar to address unmet sampling needs and challenges, which makes it possible, for the first time, to study cell differentiation and organogenesis at early embryonic stages inaccessible previously. I have mathematically validated the capability of MagCar to isolate tissue-specific cells from developing embryos and experimentally succeeded in isolating neural- and epidermal-specific cells from a wide range of embryonic stages. Using MagCar, I isolated high-purity neural- and epidermal-fated cells (92.4%) with high cell viability (96.3%) and negligible sample loss (< 0.3%). MagCar was scalable to early embryonic stages such as the gastrula stage, which is a critical stage that guides tissue formation and is challenging for existing techniques to work with because the cell sizes are too large for typical FACS instruments. Moreover, this novel strategy required no expensive and sophisticated instruments, enabling in-house sample collection from developing organisms. To demonstrate that MagCar is compatible with downstream sample processing and analysis, we developed liquid chromatography (LC)-MS methods to perform both metabolomic and proteomic analysis of the MagCar isolated neural-fated cells and identified ~80 metabolites and ~6,000 proteins, respectively. These results laid a solid foundation for the application of MagCar to study temporal molecular changes during tissue formation. Moreover, the next-generation MagCar technique can be further improved by using magnetic particles made of chemically inert and biocompatible materials, thus, minimizing molecular changes during the sample collection process. Furthermore, MagCar can potentially be applied to use functional magnetic particles to isolate subcellular compartments for studying subcellular heterogeneity. For example, magnetic particles with specific antibodies can be injected into early precursor cells to bind selected organelles during cell cleavage, which allows for tracking and isolating targeted organelles to explore the biomolecular differences at the subcellular level over embryonic development.

6.2 Future Directions

6.2.1 Explore subcellular heterogeneity

My research highlighted the importance of studying the molecular components of cells deeper to better understand their functions. The proteo-metabolomics results in my research showed that even small abundance changes can affect cellular fate decisions. Moreover, the same biomolecules play different biological functions at different subcellular locations. For example, neuropeptide Y functions as a neurotransmitter in the cytoplasm to convey signals while is involved in gene expression regulation in the nucleus.[28] In this dissertation, we developed *in vivo* microsampling to collect subcellular contents from live embryonic cells for single-cell analysis without interfering with normal embryonic development.[26] The developed sampling technique is applicable for multiple sampling of the same cells at different positions, thus, allowing exploration of subcellular molecular heterogeneity within the cell along with single-cell MS.

However, it is still challenging to achieve pinpoint microsampling of targeted subcellular compartments or organelles free of contamination. Therefore, one of the future directions for *in vivo* microsampling is to achieve better visualization of the *in vivo* microsampling process for subcellular microsampling of different cellular compartments with high purity. Combining fluorescence labeling of organelles with microscopy could be one way to visualize the sampling process of labeled organelles. For example, staining of the nucleus using Hoechst 33342 enabled accurate microsampling of the nucleus and cytoplasm of neurons, followed by single-cell MS analysis to discover the localization of peptides within cells.[28] As it comes to the era of subcellular analysis of organelles[76], we expect to see more and more studies exploring subcellular heterogeneity using the developed *in vivo* microsampling, the capillary size of which can be tailored to minimize perturbation during the sampling process for different biological systems.

6.2.2 Advances for multi-omics

The multi-omic strategy (metabolomics and proteomics) as shown in my dissertation work advanced research involved in single-cell and tissue analysis by providing complementary results in pathway analysis, which helped to reveal cell-to-cell heterogeneity at early embryonic stages[26] and to elucidate molecular mechanisms underlying metabolite-induced cell fate change. However, there is still an unmet need to measure both downstream metabolites and proteins and upstream RNAs and DNAs from a single sample, which promises to offer a more holistic picture of complex biological activities. Furthermore, it is challenging to integrate data from different ‘omic analyses to obtain meaningful information with high confidence.

Improve extraction workflows enable simultaneous isolation of metabolites, proteins, RNA, and DNA. For a systems biology understanding of the cell, all the biomolecules including transcripts, proteins, and metabolites need characterization. Recently, *in vivo* microsampling was developed to enable the sampling of the same cells in live embryos multiple times for metabolomics and proteomics.[26] This sampling technique enables multiple nondestructive sampling of the same cells from live embryos for proteo-metabolomics of single cells, which empowered us to obtain new biological knowledge. To address the sampling challenge when the sample amount is too limited to allow multiple sampling like patient samples and to remove sample-to-sample variation during multiple sampling to compare different ‘omic results[231], new sample extraction strategies and techniques were developed to enable simultaneous isolation of different types of biomolecules from the same sample for multi-omic analysis.[232-233] For example, this dissertation reports a simultaneous extraction strategy to isolate metabolites and proteins from the collected tissue-specific cells, which enabled us to conduct proteo-metabolomic analysis using LC-MS. Moreover, lipids, polar metabolites, and proteins were reported to be isolated from the same sample using an MTBE-based extraction procedure during which 2 layers of liquid phase were formed with insoluble proteins precipitated on the bottom—the top organic phase containing lipids and the bottom aqueous phase containing polar metabolites.[233] Additionally, new reagents and techniques have been developed to allow the isolation of metabolites, proteins, RNA, and DNA from the same sample.[232, 234] However, these simultaneous isolation strategies generally result in different isolation efficiency when recovering different types of biomolecules[234] which bring bias in detecting and identifying biomolecules during the sample measurement. Therefore, efforts were put into optimizing these simultaneous extraction strategies.[231, 234] Nevertheless, there is a still pressing need for unbiased simultaneous isolation of metabolites, proteins, RNA, and DNA, which will improve the extraction recovery to maximize the detection and quantification of these biomolecules, thus, advancing research in the multi-omic analysis of single cells, tissues, or organs in cell and developmental biology, neuroscience, and diseases.

Develop intelligent bioinformatics to facilitate multi-omic data integration. Due to the complexity of multi-omic datasets, it remains challenging to integrate datasets from different ‘omic analyses for meaningful biological findings. Moreover, it is imperative to identify true variations or changes and filter out false positive results during result interpretation, which requires interdisciplinary expertise. Though interlaboratory collaborations among different ‘omic backgrounds will be an effective way to address this challenge, there is an emerging need to use bioinformatics to develop software for automatic data processing, analysis, and integration. While previous efforts from the communities have developed XCMS online[13], MetaboAnalyst[224], and other software for multi-omic data integration, we are expecting to see additional software development to enable automatic data integration as different ‘omics data libraries become available. In the meantime, software based on artificial intelligence providing confident identification of key pathways or biomarkers involved in biological activities is in need.

Appendices

Appendix 3.1. Quantification of proteins and metabolites between the single L-D1 and L-V1 cells in 8-cell *X. laevis* embryos using *in vivo* subcellular HRMS. The 27 proteins and 13 metabolites driving cell type-based clustering in **Figure 3.6C** are in bold. Statistically significantly dysregulated proteins and metabolites are highlighted (paired *t*-test, $p < 0.05$). Fold change, FC, was calculated as D1/V1. Key: E#, Embryo number (e.g. E1 = Embryo #1).

Master Entry	**Gene Name**	**E1**		**E2**		**E3**		**E4**		**Mean FC**	***p*-value**
		L-D1$_1$	**L-V1$_1$**	**L-D1$_2$**	**L-V1$_2$**	**L-D1$_3$**	**L-V1$_3$**	**L-D1$_4$**	**L-V1$_4$**		
BG018089	*rpl31*	0.206	-0.087	0.110	0.401	-0.056	0.000	0.486	0.026		
BG021328	*rpl36a*	1.098	0.933	0.484	0.492	0.080	0.891	-0.204	0.626		
BG513260	*vtgb1*	-1.485	-0.291	-0.140	-0.651	2.354	2.618	2.136	2.122		
BJ056971	*rps11*	0.082	0.196	0.033	-0.622	-0.372	-0.021	0.080	-0.061		
BJ091434	*rpl11*	0.821	0.195	0.053	0.591	-0.121	-0.283	-0.135	-0.523		
BJ615339	*hist1h2bj*	0.638	0.394	0.350	0.557	0.283	-0.127	-0.016	0.212		
BM180694	*c3*	-0.673	-0.555	-1.046	-0.538	1.315	1.486	1.471	1.388		
BP687641	*rps28p9*	0.261	0.981	0.904	1.082	0.441	0.701	0.943	0.523		
BP731724	*habp4*	0.602	-0.120	-0.149	-0.096	0.072	0.057	-0.792	-0.184		
CA971511	*hbz*	1.494	1.381	1.315	1.482	1.093	0.766	1.274	0.932		
CL10048Contig1	*rplp2*	1.068	1.052	0.938	1.041	0.918	0.663	0.718	0.825		
CL1011Contig5	*vcp*	-1.541	-0.433	-0.390	-0.565	-0.079	-0.118	-0.057	-0.067		
CL10183Contig1	*eef1b2*	0.244	0.135	0.385	0.641	0.287	0.222	0.033	0.259		
CL10203Contig2	*gstt1*	-0.338	-0.230	-0.393	-0.085	0.557	0.221	0.492	0.555		
CL1026Contig6	*mapre1*	0.006	-0.211	-0.328	-0.144	-0.924	-0.139	-0.840	-0.472		
CL10303Contig1	*gstm1*	0.636	-0.184	-0.123	-0.068	0.807	0.569	0.915	0.726		
CL10311Contig1	*fabp4*	0.202	0.364	0.338	0.239	-0.531	-0.287	0.495	-0.580		
CL10318Contig1	*rpl23*	-0.213	0.124	0.379	-0.660	0.202	0.001	-0.551	0.122		
CL10546Contig1	*serpinb6*	-1.086	-0.875	-0.760	-0.950	-0.675	-0.798	-0.947	-0.885		
CL10608Contig1	*fabp4*	0.713	0.571	0.860	0.603	0.081	0.318	0.700	-0.086		

CL10772Contig1	*ccdc15*	-1.665	-2.245	-2.038	-1.529	0.157	-0.213	-0.827	-0.204
CL107Contig4	*eif5a*	-0.522	-0.076	0.029	-0.387	0.462	0.610	0.315	0.269
CL10801Contig1	*nqo1*	-0.257	-0.563	-0.570	-0.519	-2.657	-2.153	-1.214	-0.633
CL10849Contig1	*anxa4*	-0.998	-2.277	-1.053	-1.562	-1.056	-0.710	-1.558	-1.074
CL1097Contig1	*rps15a*	-0.269	0.093	0.105	-0.134	-0.055	-0.176	0.569	0.187
CL11091Contig1	*atp5j2*	-0.999	-0.773	-0.178	-0.165	-0.237	0.157	-0.861	-0.104
CL1127Contig1	*atp5b*	0.387	0.739	0.393	-0.359	0.570	0.461	1.110	0.581
CL1127Contig2	*atp5b*	0.532	0.734	0.327	-0.058	0.626	0.384	1.042	0.522
CL1144Contig2	*hadhb*	-0.146	-0.467	-0.526	-0.065	-0.829	-0.742	-0.421	-0.967
CL1144Contig4	*hadha*	-1.687	-1.266	-1.307	-0.852	-0.847	-0.588	-0.056	-0.505
CL1152Contig1	*pkm*	0.379	0.501	0.517	0.381	0.921	0.713	0.940	0.857
CL1152Contig2	*pkm*	0.273	0.319	0.337	0.327	0.774	0.498	0.692	0.647
CL116Contig7	*sod2*	-0.656	-0.538	-0.191	-0.521	-0.577	-1.456	-0.518	-0.248
CL11731Contig1	*dbi*	0.697	0.704	0.514	0.749	0.079	0.557	0.053	0.178
CL1187Contig1	*tagln2*	0.050	-0.065	0.020	0.053	-0.945	-0.494	-0.642	-0.889
CL1187Contig3	*tagln2*	0.001	-0.150	-0.020	0.063	-0.994	-0.542	-0.632	-0.937
CL11897Contig1	*fh*	-1.571	-0.727	-1.245	-1.435	-0.778	-2.371	-1.432	-1.396
CL11Contig15	*hsp90ab1*	0.129	0.276	0.068	-0.208	0.369	0.187	0.528	0.447
CL11Contig16	*hsp90ab1*	0.129	0.259	0.039	-0.218	0.368	0.158	0.524	0.445
CL1206Contig3	*ddx6*	-0.499	-1.100	-1.249	-0.526	-1.061	-0.860	-1.436	-0.876
CL12198Contig1	*eef1a2*	0.786	0.581	0.305	0.469	0.641	0.536	0.585	0.545
CL12519Contig2	*ndufb8*	0.000	-0.425	-0.965	-0.222	-1.138	-1.510	-1.271	-1.710
CL12613Contig1	*uchl1*	0.102	-0.698	-0.666	-0.147	-0.165	0.003	-0.324	-0.265
CL12647Contig1	*atpif1*	1.613	0.963	0.670	1.270	0.592	0.214	0.438	0.327
CL12806Contig1	*mgst3*	-1.497	-0.232	-0.297	-0.662	-0.049	0.037	0.373	0.155
CL12856Contig1	*ndufa13*	-0.057	0.104	-0.383	-0.244	-0.639	-0.636	-0.940	-0.392
CL134Contig2	*rpl35*	0.951	0.962	0.820	0.357	0.525	0.455	0.590	0.505
CL13623Contig1	*serpina6*	0.309	0.559	1.161	0.320	1.621	1.730	1.266	1.390
CL1365Contig1	*cap1*	-1.502	-0.668	-1.176	-0.468	-1.288	-1.345	-1.363	-1.412
CL1377Contig3	*nasp*	-1.649	-1.443	-1.323	-0.755	-0.267	-0.318	-0.117	-0.242
CL1377Contig4	*nasp*	-0.823	-0.478	-0.762	-0.392	-0.209	-0.333	-0.422	-0.205
CL137Contig1	*got2*	-0.132	0.138	0.079	-0.032	-0.438	-0.356	-0.063	-0.251
CL137Contig3	*got2*	-0.220	-0.031	-0.003	-0.027	-0.348	-0.343	-0.099	-0.283

CL13915Contig1	*cd63*	-1.071	-0.920	-0.745	-0.936	-0.588	-0.114	-0.933	-0.715		
CL13Contig12	*ncl*	-1.663	-0.938	-1.337	-0.687	-1.123	-2.463	-1.524	-1.337		
CL13Contig16	*eif3a*	-0.448	-1.545	-1.294	-0.743	-1.367	-1.214	-1.525	-1.695		
CL1402Contig1	*tuba4a*	-0.678	0.256	0.204	0.001	0.817	0.696	0.890	0.721		
CL1402Contig2	*tuba1a*	-0.614	0.459	0.365	0.221	0.896	0.735	0.799	0.707		
CL1402Contig3	*tuba1b*	-0.678	0.248	0.167	0.001	0.921	0.748	0.985	0.787		
CL14087Contig1	*c21orf33*	-0.710	-1.290	-0.384	-0.575	-0.144	-0.158	-0.571	-0.253		
CL14502Contig1	*tstd3*	-1.113	-0.511	-0.787	-0.143	-0.994	-1.913	-0.974	-2.398		
CL14883Contig1	*rplp2*	0.677	0.602	0.396	0.571	-0.909	-0.630	-0.391	-0.383		
CL148Contig10	*rps2*	0.818	0.563	0.405	0.535	0.491	0.376	0.585	0.417		
CL148Contig11	*rps2*	0.818	0.584	0.438	0.522	0.456	0.371	0.665	0.425		
CL1514Contig2	***eno1***	**1.127**	**1.089**	**1.230**	**0.968**	**1.277**	**1.077**	**1.423**	**1.141**	**1.61**	**0.039**
CL1515Contig2	*cct8*	-1.272	-0.574	-0.938	-0.841	0.061	-0.132	0.017	-0.102		
CL15183Contig1	*hbz*	1.480	1.382	1.314	1.406	1.174	0.965	1.308	1.029		
CL152Contig10	*eef1d*	0.555	0.405	0.175	0.594	0.257	0.086	0.376	0.138		
CL1548Contig2	*rpl5*	0.592	0.709	0.493	-0.022	0.251	0.179	0.231	0.191		
CL1548Contig4	*rpl5*	0.430	0.429	0.221	-0.262	0.062	0.028	0.269	-0.084		
CL15613Contig1	***uchl3***	**-0.867**	**-0.713**	**-0.879**	**-0.208**	**-0.723**	**-0.488**	**-1.427**	**-0.193**	**0.39**	**0.103**
CL1582Contig2	*rpl13a*	0.403	0.264	0.202	-0.061	-0.022	-0.189	-0.179	-0.252		
CL15Contig35	*anxa7*	-0.467	-1.034	-0.632	-0.488	-1.026	-2.122	-1.183	-0.386		
CL1612Contig3	*phb*	-0.134	-0.031	-0.387	-0.164	-0.062	-0.409	-0.859	-0.311		
CL1612Contig4	*phb*	0.078	0.101	-0.162	0.075	0.093	0.021	-0.621	-0.014		
CL1627Contig2	*aldoa*	0.961	0.818	0.711	0.538	1.145	0.917	0.968	0.958		
CL1631Contig2	*rpl34*	1.225	0.303	0.400	0.842	0.011	0.157	0.146	-0.218		
CL1634Contig2	*pa2g4*	-0.500	-0.384	-0.454	-0.259	-0.055	0.032	-0.167	-0.188		
CL1634Contig3	*pa2g4*	-0.500	-0.156	-0.454	-0.259	0.043	0.098	-0.167	0.747		
CL1655Contig4	*rpl29*	1.292	0.145	0.172	0.723	0.123	0.480	0.181	0.066		
CL1665Contig1	*ywhae*	-0.727	-0.720	-0.593	-0.707	-0.489	-0.673	-0.829	-0.390		
CL1684Contig4	*fkbp2*	-0.734	-0.616	-0.139	0.252	-0.383	-0.145	-0.596	-0.352		
CL17153Contig1	***cox5b***	**-0.240**	**0.400**	**-0.089**	**0.782**	**-2.533**	**-0.189**	**-0.390**	**-0.366**	**0.33**	**0.143**
CL1744Contig1	*pdia3*	0.543	0.322	0.295	0.284	-0.088	-0.176	-0.117	-0.391		
CL1744Contig2	*pdia3*	0.229	0.164	0.106	0.112	0.068	-0.220	-0.094	-0.466		
CL1744Contig3	*pdia3*	0.786	0.500	0.487	0.796	0.116	0.234	0.384	-0.036		

CL17705Contig1	*rpsa*	0.393	0.781	0.745	0.496	0.503	0.317	0.214	0.609		
CL17837Contig1	*phpt1*	-0.067	-0.371	-0.583	0.057	-2.214	-1.709	-0.771	-2.195		
CL17983Contig1	*xb5727280*	-0.772	-1.049	-1.146	-0.328	-0.236	-0.474	-0.014	-0.580		
CL18403Contig1	*aldoa*	0.652	0.264	0.184	0.245	0.820	0.545	0.364	0.570		
CL18578Contig1	*uchl1*	0.029	-1.647	-0.692	-0.232	-1.062	-1.866	-0.928	-0.763		
CL1892Contig2	*st13*	-0.436	-0.292	-0.434	0.012	-2.440	-1.935	-0.997	-2.421		
CL19274Contig1	*apob*	1.963	0.910	1.743	1.445	3.023	3.154	2.895	2.832		
CL1948Contig3	*cct5*	-1.886	-0.681	-0.821	-1.051	-0.001	-0.422	-0.799	-0.160		
CL1948Contig4	*cct5*	-1.896	-0.624	-1.025	-1.062	-0.094	-0.453	-0.810	-0.272		
CL19654Contig1	*pkm*	0.088	-0.293	-0.298	-0.335	0.034	-0.171	-0.123	0.052		
CL19743Contig1	*hspa8*	-0.249	0.254	0.131	-0.218	0.486	0.031	0.245	0.223		
CL1995Contig4	*ak2*	-0.341	-0.250	-0.669	-0.066	0.574	0.969	-0.260	0.252		
CL1Contig1214	*rps27a*	0.907	0.789	0.727	0.696	-0.463	-0.189	0.001	-0.602		
CL1Contig1336	*rpl17*	0.516	0.882	0.793	0.521	0.329	0.401	0.754	0.450		
CL1Contig1406	***ckb***	**0.554**	**0.387**	**0.458**	**0.374**	**0.533**	**0.414**	**0.737**	**0.552**	**1.60**	**0.039**
CL1Contig1549	*eif4a2*	-0.104	-0.465	-0.737	-0.572	-0.445	-0.299	-0.294	-0.495		
CL1Contig271	*anxa5*	-1.060	-0.866	-0.734	-0.925	-0.249	-0.228	0.000	-0.151		
CL1Contig276	*rpl28*	1.063	0.528	0.559	0.776	0.143	-0.060	0.080	0.007		
CL1Contig277	*ranbp1*	0.572	0.182	0.117	0.737	0.001	0.073	-0.168	-0.041		
CL1Contig313	*ctsc*	-1.270	-0.715	-0.944	-1.135	-0.771	-0.081	-0.252	-0.364		
CL1Contig372	*rps17*	0.769	0.483	0.601	0.820	0.828	0.529	0.644	0.689		
CL1Contig514	*ckb*	0.963	0.761	0.764	0.905	0.767	0.527	0.711	0.661		
CL1Contig529	*rpl10*	0.470	0.353	0.393	0.549	0.046	-0.009	0.208	-0.059		
CL1Contig546	*rpl6*	1.192	0.806	0.659	0.863	0.300	0.204	0.590	0.328		
CL1Contig724	*ndufa7*	0.439	0.059	-0.327	0.344	-0.778	-0.802	-0.563	-0.806		
CL1Contig763	*rpl6*	1.181	0.802	0.658	0.843	0.286	0.204	0.590	0.293		
CL1Contig796	*rpl10*	0.558	0.392	0.497	0.612	0.188	-0.031	0.275	0.061		
CL1Contig80	*c21orf33*	-0.061	-0.030	-0.132	0.120	-0.449	-1.874	-0.236	0.059		
CL1Contig853	*hist1h4a*	-0.148	0.074	0.190	0.000	-0.005	0.298	-0.708	-0.224		
CL1Contig861	*ranbp1*	0.604	0.214	0.112	0.583	-0.048	-0.105	-0.136	-0.137		
CL1Contig980	*myh10*	-2.338	-1.874	-2.012	-2.202	-0.964	-1.280	-2.199	-1.837		
CL2004Contig2	*sri*	0.125	0.485	0.399	0.176	-0.135	0.256	-0.520	-0.371		
CL2004Contig4	*sri*	0.063	0.293	0.242	-0.018	-0.197	0.193	-0.714	-0.433		

CL20154Contig1	***eno1***	**0.900**	**1.010**	**1.112**	**0.832**	**1.100**	**0.871**	**1.224**	**0.917**	**1.61**	**0.039**
CL2028Contig2	*hspa5*	-0.092	0.115	0.290	0.230	-1.245	-0.081	-0.444	-0.654		
CL2090Contig3	*rrm1*	-1.935	-1.498	-1.609	-1.800	-0.586	-0.953	-1.797	-0.671		
CL2097Contig2	*ahcy*	-0.521	-0.152	-0.041	0.041	0.514	0.174	0.229	0.359		
CL2106Contig4	*fkbp10*	-0.537	-0.824	-1.198	-0.474	-2.828	-2.324	-1.385	-2.809		
CL21119Contig1	*eef2*	-0.826	-0.065	-0.312	0.252	0.149	0.307	-0.003	0.151	**0.46**	**0.074**
CL2114Contig1	*atp5h*	0.121	-0.174	-0.569	0.599	-0.692	-0.229	0.017	-0.453		
CL2114Contig3	*atp5h*	0.383	0.023	-0.372	0.657	-2.283	-1.079	0.017	-1.198		
CL2126Contig4	*fscn1*	-1.577	-1.352	-1.214	-0.743	-1.162	-1.034	-1.438	-0.938	**0.50**	**0.036**
CL21870Contig1	*avd2*	0.291	-0.030	0.599	-0.149	-0.366	-0.061	-0.845	-0.436		
CL21967Contig1	*tmsb10*	0.557	-0.983	0.014	0.549	-2.407	-1.902	-0.963	-2.388		
CL2211Contig2	***gsn***	**-1.578**	**-0.625**	**-0.623**	**-0.744**	**-0.715**	**-0.642**	**-0.477**	**-0.584**	**0.49**	**0.160**
CL2211Contig3	*gsn*	-1.599	-0.351	-0.725	-0.764	-1.117	-0.893	-1.460	-0.803		
CL22Contig15	*pygl*	-0.470	-0.926	-0.718	-1.267	0.051	-0.198	0.143	0.093		
CL22Contig18	*rpl32*	1.217	0.863	0.850	0.950	0.408	0.357	0.345	0.287		
CL22Contig4	*rps20*	0.222	1.213	0.772	0.675	0.558	0.546	0.901	0.517		
CL22Contig8	*rps20*	0.087	1.263	0.809	0.730	0.645	0.599	1.091	0.652		
CL2318Contig2	*rpl15*	1.210	0.852	0.635	0.537	0.457	0.079	0.555	0.496		
CL2319Contig2	*ppib*	0.107	-0.173	0.108	-0.079	-0.409	-0.152	-0.775	-0.550		
CL2323Contig3	*rps6*	0.814	0.390	0.290	0.402	-0.015	-0.046	0.396	-0.026		
CL2400Contig2	*ldhb*	0.566	0.521	0.535	0.374	0.955	0.550	1.241	0.794		
CL24037Contig1	*hsp90ab1*	-1.395	-1.276	-0.885	-1.259	0.196	0.495	0.449	0.248		
CL2410Contig2	*rps5*	0.503	0.530	0.421	0.003	0.120	-0.042	0.645	0.420		
CL2418Contig1	*cct4*	-1.055	-0.842	-0.729	-0.920	-0.237	-0.457	-0.916	-0.170		
CL2418Contig2	*cct4*	-1.078	-0.793	-0.752	-0.943	-0.218	-0.475	-0.940	-0.192		
CL247Contig3	*loc392647*	-0.508	-1.072	-0.618	0.147	-0.803	-2.007	-1.068	-2.492		
CL25251Contig1	*tmsb4x*	0.269	0.496	0.415	1.084	-1.735	-1.230	-0.292	-1.716		
CL2555Contig3	*gdi2*	-0.704	-0.825	-0.898	-0.910	-0.696	-0.854	-1.605	-0.871		
CL2615Contig2	*atp5a1*	0.524	0.319	-0.073	-0.073	0.376	0.316	0.857	0.264		
CL2615Contig3	*atp5a1*	0.532	0.249	-0.156	-0.190	0.370	0.293	0.863	0.274		
CL26379Contig1	*actb*	1.683	1.417	1.398	1.512	1.380	1.328	1.592	1.243		
CL26721Contig1	*tubb4b*	-0.928	-0.498	-0.651	-0.734	0.152	-0.012	0.295	-0.142		
CL2685Contig1	***tpi1***	**0.975**	**1.047**	**1.071**	**0.767**	**0.943**	**0.857**	**1.086**	**0.809**	**2.07**	**0.119**

CL2691Contig4	*rpia*	-1.084	-0.479	-0.283	0.212	-2.388	-1.884	-0.246	-1.156		
CL2807Contig6	*acadm*	-0.805	-0.247	-0.417	-0.612	-0.582	-0.448	-1.366	-0.845		
CL2871Contig3	*lmf2*	-1.408	-1.988	0.580	0.104	-2.013	0.629	-1.269	0.144		
CL289Contig3	*hsp90b1*	-1.643	-1.244	-1.317	-1.507	-0.622	-0.248	-0.213	-0.660		
CL289Contig7	*rps14*	1.488	1.180	1.082	1.298	0.718	0.563	0.673	0.599		
CL2962Contig3	*rps16*	-0.185	0.703	0.485	0.287	-0.086	0.018	0.298	-0.195		
CL2985Contig3	*txndc5*	-1.219	-0.884	-0.554	-0.385	-1.214	-2.019	-1.081	-2.505		
CL2995Contig1	*cct2*	-0.554	-0.794	-0.610	-1.118	0.512	0.184	0.140	0.333		
CL2995Contig3	*cct2*	-0.564	-0.912	-0.732	-1.128	0.397	0.124	0.040	0.190		
CL29Contig3	*slc25a5*	-1.197	-0.675	-1.159	0.024	-0.060	0.394	0.031	0.022		
CL2Contig33	***hagh***	**-0.432**	**-0.763**	**-1.300**	**-0.373**	**-2.930**	**-0.299**	**-1.487**	**-0.785**	**0.62**	**0.208**
CL303Contig2	*flna*	-0.939	-1.012	-1.065	-0.589	-2.942	-2.438	-1.499	-2.924		
CL303Contig4	*flna*	-1.168	-1.430	-1.376	-0.976	-3.171	-2.667	-1.728	-3.152		
CL3082Contig1	*tcp1*	-0.854	-0.573	-0.573	-0.977	-0.059	-0.259	-0.082	-0.177		
CL3171Contig1	*rpl7*	1.353	0.934	0.817	1.137	0.717	0.690	0.706	0.722		
CL3171Contig2	*rpl7*	1.096	0.765	0.719	0.981	1.014	0.805	0.521	0.864		
CL3303Contig2	*cct7*	-2.078	-0.490	-0.735	-1.244	0.034	-0.258	-0.049	0.043		
CL3329Contig1	*pgk1*	0.074	-0.143	0.232	-0.986	0.651	0.640	0.816	0.614		
CL3329Contig2	*pgk1*	0.047	-0.161	0.201	-0.873	0.676	0.651	0.782	0.645		
CL3418Contig2	*gapdh*	1.535	1.152	1.113	1.368	1.225	1.093	1.398	1.179		
CL3430Contig1	*psma1*	-0.175	-0.368	-0.738	-0.516	-0.544	-0.235	-1.212	-0.874		
CL3430Contig2	*psma1*	0.088	-0.404	-0.552	-0.193	-0.488	-0.180	-0.888	-0.738		
CL3441Contig3	*hibadh*	-1.735	-1.115	-0.893	-0.900	-0.759	-2.535	-1.596	-0.749		
CL3453Contig2	*akr1b1*	-0.659	-0.579	-0.363	-1.222	-0.347	0.230	-0.322	-0.239		
CL3453Contig3	*akr1b1*	-0.695	-0.615	-0.362	-1.259	-0.727	-0.318	-0.358	-0.727		
CL3505Contig1	*aldoc*	0.662	0.544	0.496	0.732	0.722	0.457	0.563	0.663		
CL3506Contig1	*park7*	0.511	0.226	0.327	0.555	0.270	0.394	-0.081	0.152		
CL3506Contig2	*park7*	0.613	0.322	0.392	0.799	0.280	0.386	-0.110	0.204		
CL3526Contig1	*ppa1*	-0.475	-1.054	-1.257	-0.787	-2.925	-2.421	-1.482	-2.907		
CL36281Contig1	*frg1*	0.341	-1.185	-0.977	0.230	-0.452	0.229	-0.466	-0.257		
CL3649Contig1	*cyb5r3*	-1.110	-0.806	-0.359	-0.970	0.001	0.000	-0.268	-0.238		
CL3652Contig1	*prdx6*	-0.256	0.140	0.388	0.003	0.327	0.123	0.622	0.375		
CL3654Contig2	*hspa9*	-0.229	-0.473	-0.566	-0.458	-1.145	-0.892	-1.154	-0.986		

CL367Contig6	*eif4a1*	-0.153	-0.459	-0.754	-0.621	-0.261	-0.256	-0.344	-0.246		
CL3695Contig1	*pfn1*	-0.955	-0.446	-0.254	-0.121	-0.102	-0.348	0.030	-0.379		
CL3721Contig3	*psmd2*	-1.189	-1.369	-1.478	-1.753	-1.011	-1.134	-1.749	-1.119		
CL3748Contig2	*cbr1*	-1.068	-0.555	-0.021	0.051	-0.998	-1.169	-0.929	-1.107		
CL384Contig8	*rpl11*	0.035	0.782	0.515	0.443	0.305	0.161	0.370	0.341		
CL3868Contig2	*rplp1*	0.225	-0.079	0.622	1.245	0.572	0.464	0.314	0.685		
CL39056Contig1	*ybx2*	0.622	0.396	0.087	0.611	-0.384	0.039	-0.161	0.111		
CL3918Contig2	*cndp2*	-0.644	-0.382	-0.120	-0.330	-0.081	-0.339	-0.106	-0.145		
CL391Contig4	*rpl18a*	0.692	0.238	0.176	0.280	0.225	0.108	0.490	0.258		
CL3941Contig2	*fam115c*	-0.835	-1.014	-1.019	-1.115	-0.653	-1.232	-1.306	-0.843		
CL394Contig6	*eef1g*	0.844	0.491	0.439	0.874	0.097	-0.051	0.333	0.150		
CL40282Contig1	*hist2h2ab*	0.549	0.024	-0.028	0.559	-1.659	-1.155	-0.216	-0.397		
CL4086Contig2	*pgam1*	0.186	0.031	0.300	0.633	0.051	0.048	-0.415	0.182		
CL40Contig4	*plin2*	0.200	-0.188	0.824	0.391	0.851	0.996	0.692	0.514		
CL4112Contig2	*myl6*	-0.010	-0.397	-0.338	-0.330	-0.737	-1.965	-1.026	-0.774		
CL4166Contig1	*stip1*	-1.944	-1.063	-1.011	-1.110	-2.034	-2.744	-1.806	-3.230		
CL4254Contig1	*tubb3*	-1.246	0.356	0.341	-0.412	0.772	0.641	0.863	0.779		
CL4254Contig2	*tubb3*	-1.246	0.356	0.341	-0.412	0.777	0.645	0.863	0.787		
CL4267Contig1	*tkt*	-0.263	-0.001	-0.034	-0.008	0.579	0.358	0.642	0.473		
CL4286Contig1	*itln2*	-1.028	-0.409	0.032	-0.193	0.289	0.546	0.260	0.244		
CL4288Contig3	*rpl24*	1.247	0.794	0.779	0.919	0.579	0.426	0.589	0.473		
CL4312Contig2	*rpl23a*	0.640	0.610	0.780	1.179	0.473	0.447	0.902	0.665		
CL4343Contig2	*gyg1*	-1.270	-0.695	-0.321	-0.436	-1.172	-1.150	-1.131	-1.263		
CL437Contig2	*p4hb*	-0.486	-0.563	-0.531	-0.455	-0.625	-0.736	-1.150	-0.461		
CL4398Contig3	*ndufb3*	-0.098	-0.288	-0.523	-0.477	-2.615	-0.437	-1.172	-0.922		
CL4491Contig1	*mdh2*	0.823	1.111	0.988	1.146	0.660	0.477	1.018	0.584		
CL4491Contig2	*mdh2*	0.766	1.079	0.939	1.047	0.635	0.442	0.996	0.523		
CL45078Contig1	*rpl30*	1.340	1.142	0.968	1.429	0.238	-1.001	0.636	0.337		
CL4522Contig1	***atp5f1***	**-0.117**	**-0.374**	**-0.441**	**-0.927**	**-0.655**	**-0.580**	**-0.224**	**-0.634**	**1.78**	**0.095**
CL4522Contig2	*atp5f1*	-0.139	-0.198	-0.422	-0.813	-0.677	-0.602	-0.110	-0.656		
CL4552Contig2	*aloxe3*	-1.018	-1.450	-1.674	-1.865	-0.041	-0.229	0.079	-0.011		
CL45Contig20	***naca***	**0.339**	**0.448**	**0.308**	**0.702**	**0.116**	**0.229**	**0.055**	**0.344**	**0.62**	**0.048**
CL45Contig6	*naca*	0.683	0.452	0.366	0.598	0.134	0.235	0.055	0.378		

CL4718Contig1	*hbz*	1.473	1.803	1.760	1.715	1.446	1.279	1.711	1.409		
CL47Contig4	*myh9*	-1.202	-1.062	-1.278	-1.281	-1.601	-1.244	-1.037	-1.178		
CL4803Contig4	*psmd7*	-1.567	-0.573	-1.241	-1.431	-1.333	-2.367	-1.428	-0.981		
CL4816Contig2	*etfa*	-1.194	-0.926	-0.809	-0.360	-0.994	-1.994	-1.056	-0.775		
CL4821Contig2	*paics*	-0.945	-0.757	-0.680	-0.554	-0.234	-0.085	-1.505	-0.423		
CL48Contig2	*rps27*	0.792	0.505	0.370	0.408	0.284	0.408	0.088	0.138		
CL48Contig6	*rps27*	0.976	0.692	0.520	0.653	0.471	0.552	0.167	0.375		
CL4966Contig1	*suclg1*	-0.445	-0.192	-0.402	-0.326	-0.342	0.110	-0.052	-0.394		
CL4966Contig2	*suclg1*	-0.470	-0.217	-0.426	-0.351	-0.367	0.085	-0.077	-0.419		
CL5043Contig1	*npm2*	-0.527	-0.408	0.212	0.315	-0.306	-1.327	-0.388	-0.915		
CL50Contig6	*rps11*	1.152	0.716	0.576	0.738	0.171	0.253	0.464	0.307		
CL5138Contig2	*pkm*	-1.030	-0.508	-0.569	-0.895	-0.444	-0.485	-0.193	-0.435		
CL51Contig4	*cpne1*	-2.200	-0.767	-1.014	-1.365	-1.466	-3.000	-2.061	-1.197		
CL52Contig20	*hspa8*	0.097	0.402	0.308	0.157	0.573	0.156	0.208	0.346		
CL52Contig8	*hspa8*	0.115	0.390	0.289	0.146	0.586	0.154	0.235	0.352		
CL5313Contig1	*rpl9*	1.443	-0.049	0.057	-0.035	-0.018	0.235	0.373	0.082		
CL5360Contig1	*vwa5a*	-1.451	-1.333	-0.931	-1.316	-0.518	-0.612	-1.313	-0.517		
CL5376Contig1	*dars*	-1.711	-1.442	-1.183	-2.274	-0.389	-0.388	-2.271	-0.754		
CL53Contig3	*tkt*	-1.068	-1.036	-0.829	-0.964	-1.479	-1.294	-1.659	-1.685		
CL5407Contig1	*gsto1*	0.085	0.259	0.274	-0.137	0.107	0.412	0.099	0.050		
CL5428Contig3	***psmb2***	**-1.083**	**-0.069**	**-0.757**	**0.153**	**-0.280**	**0.178**	**-0.945**	**-0.452**	**0.22**	**0.015**
CL5505Contig1	*grhpr*	0.150	0.515	0.796	0.527	0.798	0.558	0.871	0.719		
CL5527Contig2	*cox5a*	-0.040	0.025	-0.143	0.583	-0.260	-0.370	-0.176	-0.397		
CL561Contig1	*rpl12*	1.040	0.274	0.532	0.618	0.448	0.115	0.380	0.333		
CL561Contig9	*rpl12*	0.991	0.078	0.403	0.382	0.488	0.160	-0.128	0.232		
CL569Contig8	*rps29*	1.306	0.671	0.667	0.893	0.489	0.218	0.681	0.401		
CL575Contig2	***rpl30***	**-0.775**	**0.359**	**0.078**	**0.059**	**-0.134**	**0.030**	**-0.637**	**0.127**	**0.49**	**0.151**
CL5790Contig2	*rps18*	-0.350	0.809	0.632	0.328	-0.555	-0.401	0.358	-0.041		
CL5827Contig1	*rps23*	1.040	0.826	0.673	0.568	0.676	0.391	0.913	0.562	**1.73**	**0.135**
CL5832Contig1	*gpi*	0.040	0.491	0.409	0.196	0.886	0.578	0.692	0.783		
CL5832Contig2	*gpi*	-0.149	0.306	0.178	0.004	0.648	0.295	0.385	0.414		
CL593Contig1	***ywhaq***	**-0.349**	**-0.050**	**-1.421**	**-0.152**	**-0.401**	**-0.431**	**-0.909**	**-0.558**	**0.52**	**0.189**
CL5948Contig1	*rpl39*	1.181	1.189	0.977	1.396	0.876	1.013	1.307	0.991		

CL5962Contig1	*taldo1*	-0.856	-0.428	-0.307	-1.420	-0.156	-0.491	-1.416	-0.202		
CL6048Contig1	*psma4*	-1.026	-1.606	-1.399	-0.891	-0.233	-0.443	-0.188	-0.423		
CL60Contig6	*nme2*	0.834	0.952	0.945	0.966	0.949	0.852	1.295	0.903		
CL627Contig2	*acaa2*	-1.341	-0.115	-0.407	-0.507	-1.152	-0.905	-1.203	-1.115		
CL6336Contig2	*ndufs3*	0.000	-0.046	-0.231	0.049	-2.090	-1.585	-0.647	-2.071		
CL640Contig5	*aco2*	-0.216	-0.196	-0.181	-0.249	-0.582	-0.659	-0.150	-0.728		
CL640Contig7	*aco2*	-0.216	-0.087	-0.118	-0.135	-0.588	-0.549	-0.150	-0.827		
CL6511Contig1	*lgals3*	0.638	0.310	0.264	0.513	-0.097	-0.395	-0.182	-0.298		
CL6531Contig1	*tyrp1*	-0.205	-0.800	-0.568	-0.797	-2.936	-2.431	-1.493	-2.917		
CL654Contig1	*hspe1*	1.064	0.977	0.791	1.266	1.155	0.077	0.203	-0.522		
CL654Contig2	*hspd1*	0.171	0.018	0.023	0.156	0.006	0.211	0.366	0.140		
CL654Contig4	*hspd1*	0.189	0.062	0.086	0.156	0.021	0.255	0.446	0.218		
CL665Contig7	*rpl3*	0.558	0.405	0.302	0.248	0.156	-0.040	0.267	0.330		
CL6675Contig1	*acadl*	0.172	-0.128	-0.476	-0.429	-2.844	-0.912	-0.702	-1.144		
CL6681Contig1	*vdac2*	-1.208	-0.349	-0.882	-1.073	-0.143	-0.111	-0.356	0.037		
CL66Contig6	*cct6a*	-0.578	-0.379	-0.330	-0.381	0.242	-0.184	-0.277	0.038		
CL6751Contig2	*atic*	-0.897	-0.034	-0.596	-0.403	-0.049	-0.281	-0.176	0.031		
CL6751Contig3	*atic*	-0.897	-0.034	-0.596	-0.403	-0.057	-0.283	-0.176	0.016		
CL678Contig5	*uba1*	-1.799	-1.410	-1.473	-1.664	-1.186	-0.854	-0.826	-0.760		
CL6842Contig1	*habp4*	0.651	0.262	0.072	0.498	0.303	0.110	0.276	0.291		
CL6995Contig1	***cox4i2***	**-0.157**	**0.155**	**0.048**	**0.351**	**-0.343**	**-0.266**	**-0.717**	**-0.140**	**0.52**	**0.053**
CL7150Contig1	*anxa1*	-1.776	-2.357	-1.450	-1.641	-0.442	-0.359	-1.638	-1.081		
CL7196Contig1	*got1*	-0.722	-0.490	-0.115	-0.294	-0.740	-0.764	-1.282	-0.973		
CL7426Contig1	*atp5j*	0.264	-0.158	-0.543	0.169	-2.173	-1.669	-0.730	-2.155		
CL749Contig1	*btf3l4*	-1.009	-0.407	-0.190	-0.874	-1.022	-1.809	-0.172	-0.324		
CL7642Contig1	*gdi1*	-1.692	-2.272	-1.366	-1.557	-1.806	-1.451	-1.553	-1.458		
CL7665Contig1	*hadh*	-0.658	-0.255	-0.444	-1.221	-0.875	-0.306	-0.334	-0.545		
CL7695Contig1	*aldh2*	-1.487	-0.701	-1.161	-1.352	-0.857	-2.287	-1.348	-0.764		
CL76Contig9	*rps8*	1.251	0.884	0.845	1.009	0.669	0.375	0.713	0.629		
CL7800Contig2	*rpl26*	1.107	0.783	0.654	0.777	0.449	0.282	0.398	0.426		
CL7950Contig1	*gapdh*	1.562	1.226	1.143	1.392	1.273	1.092	1.433	1.109		
CL7997Contig2	***shmt1***	**-1.641**	**-0.803**	**-0.990**	**-0.806**	**-1.372**	**-0.763**	**-1.502**	**-0.993**	**0.34**	**0.029**
CL801Contig1	*eef2*	0.489	0.343	0.089	0.304	0.517	0.265	0.407	0.465		

CL801Contig2	*eef2*	0.385	0.306	0.061	0.499	0.472	0.161	0.346	0.432		
CL8065Contig1	*gsta1*	0.064	-0.916	-0.416	0.123	-2.339	-1.835	-0.896	-2.320		
CL8144Contig1	*hpgds*	-0.975	0.328	0.254	0.327	0.636	0.639	0.939	0.634		
CL8227Contig1	*rplp2*	0.791	0.707	0.649	0.866	0.456	0.060	0.602	0.296		
CL8227Contig2	*rplp2*	0.798	0.659	0.626	0.900	0.449	0.071	0.349	0.241		
CL8308Contig2	*rpl37*	0.946	0.343	0.112	0.530	0.139	-0.261	-0.166	0.094		
CL8453Contig1	*lap3*	-1.390	-0.420	-1.064	-0.363	-0.027	-0.008	-1.251	-0.270	**0.34**	**0.060**
CL8454Contig1	*habp4*	0.613	0.198	-0.038	0.465	-0.183	0.064	-0.417	0.025		
CL8614Contig1	*sod1*	-0.550	-0.584	-0.857	-0.277	0.408	0.483	0.653	0.308		
CL8653Contig1	*itln2*	-0.989	-0.450	-0.035	-0.154	0.279	0.530	0.219	0.215		
CL8746Contig1	*atp5d*	-0.064	-0.047	0.073	-0.205	-0.105	-0.080	-0.901	-0.336		
CL8757Contig1	***cox6c***	**-0.426**	**-0.157**	**-0.505**	**-0.209**	**-0.146**	**-0.136**	**-0.336**	**-0.090**	**0.65**	**0.052**
CL8833Contig1	*dhrs11*	-1.706	-2.287	-1.209	-0.872	-1.050	-2.507	-1.568	-0.996		
CL8Contig14	*ppif*	1.090	1.050	1.095	1.383	0.553	0.425	0.903	0.414		
CL9021Contig1	*ctbs*	-1.166	-0.834	-0.304	-0.331	-0.519	-1.966	-1.027	-0.942		
CL9064Contig1	*ptgr1*	-1.340	-1.222	-0.834	-1.205	-0.129	-0.183	-0.277	-0.285		
CL907Contig5	*atp6v1a*	-1.622	-1.362	-1.296	-1.487	-1.348	-2.422	-1.484	-1.444		
CL912Contig5	*rpsa*	0.476	0.856	0.795	0.672	0.455	0.294	0.188	0.579		
CL9216Contig2	*pgm1*	-0.796	-0.855	-0.661	-1.082	-0.787	-1.111	-1.778	-0.719		
CL926Contig3	*rpl18*	0.774	0.466	0.405	0.684	0.446	0.431	0.398	0.457		
CL929Contig2	*rpl38*	0.783	1.095	0.792	0.310	0.605	0.297	0.971	0.464		
CL92Contig12	*rps3*	0.397	0.788	0.639	0.480	0.590	0.460	0.817	0.536		
CL92Contig7	*rpl4a*	0.970	0.750	0.517	0.643	0.401	0.294	0.329	0.387		
CL9522Contig1	*eci1*	-0.619	-0.834	-0.643	-1.044	-3.182	-2.678	-1.739	-3.163		
CL964Contig2	*gnb2l1*	-0.252	-1.160	-1.651	-1.143	0.485	0.538	0.399	0.430		
CL9731Contig2	*cox7c*	0.126	0.024	-0.095	-0.221	-0.565	-1.856	-0.917	-0.007		
CL987Contig4	*tpm3*	-0.280	-0.954	-0.734	-0.529	-2.668	-2.163	-1.225	-2.649		
CL990Contig7	*pck2*	-2.159	-1.465	-0.933	-1.325	-0.565	-2.959	-2.021	-3.444		
CV075495	*aldh6a1*	-0.969	-0.549	-0.643	-0.834	-0.704	-0.664	-0.831	-0.820		
DC000291	*hint1*	-0.076	0.485	0.091	0.247	-2.079	-1.575	-0.636	-2.061		
DC015876	*rpl35a*	-0.055	0.002	-0.143	-0.045	-0.186	-0.065	0.359	0.390		
DC044200	*hist1h4a*	0.073	0.368	0.209	0.173	-0.435	-0.066	-0.522	-0.407		
DC088876	*prdx2*	0.119	0.301	0.384	-0.228	0.503	0.597	0.136	0.369		

dsrrswapns_comp203245_c0_seq1	*thap4*	0.165	0.069	-0.181	-0.105	-2.244	-1.740	-0.801	-2.225		
dsrrswapns_comp204545_c0_seq1	*mgst3*	-0.857	-0.218	-0.531	-0.722	-0.133	-0.023	0.240	0.051		
dsrrswapns_comp205895_c1_seq1	*pdia4*	-0.858	-0.959	-0.920	-0.645	-2.861	-2.357	-1.418	-2.842		
dsrrswapns_comp207672_c0_seq1	*aars*	-1.685	-2.265	-1.359	-1.549	-1.098	-0.870	-0.772	-1.093		
dsrrswapns_comp207966_c0_seq3	*ftcd*	-0.732	-0.882	-0.825	-0.509	-1.449	-1.528	-1.292	-1.474		
dsrrswapns_comp208676_c3_seq2	*rplp0*	-0.285	0.001	-0.043	-0.321	0.094	0.340	0.440	0.417		
dsrrswapns_comp209966_c0_seq1	*cndp2*	-1.527	-0.424	-0.230	-0.692	-0.093	-0.236	-0.119	-0.086		
dsrrswapns_comp210383_c5_seq1	*tsta3*	-1.494	-1.040	-1.168	-1.359	-0.541	-0.677	-1.356	-0.660		
dsrrswapns_comp210604_c0_seq5	*plin2*	-0.414	-0.206	0.665	0.151	0.770	0.903	0.681	0.453		
dsrrswapns_comp211382_c2_seq4	*ywhaq*	-0.365	-0.192	-0.718	-0.130	-0.275	-0.409	-0.925	-0.421	**0.65**	**0.186**
dsrrswapns_comp211955_c0_seq2	*ywhaz*	-1.218	-0.567	-0.892	-0.250	-0.462	-0.643	-0.116	-0.497		
dsrrswapns_comp212198_c1_seq14	*anxa1*	-0.635	-0.340	-0.667	-0.312	-2.638	-0.900	-1.195	-1.430		
dsrrswapns_comp212502_c2_seq1	*unknown*	-1.536	-1.097	-1.210	-1.401	-0.020	-0.243	-0.471	0.092		
dsrrswapns_comp212612_c0_seq7	*krt8*	-1.140	-0.927	-0.814	-1.005	-0.937	-0.144	-1.001	-0.671		
dsrrswapns_comp212871_c1_seq1	*mdh1*	-1.443	0.011	-0.499	-0.608	0.685	0.546	0.671	0.749		
dsrrswapns_comp213050_c0_seq18	*mlec*	-0.031	-0.055	-0.712	-0.062	-2.343	-1.838	-0.900	-2.324		
dsrrswapns_comp213410_c2_seq1	*cct6a*	-0.690	-0.491	-0.471	-0.494	0.118	-0.286	-0.303	-0.056		
dsrrswapns_comp216613_c0_seq1	*pc*	-1.775	-0.822	-2.148	-1.640	-1.449	-0.990	-0.937	-1.151		
EB470951	*tuba3c*	-1.155	-1.735	-0.829	-1.019	0.520	0.623	-1.016	0.086		
sp|A3KMT2|DAL1A_XENLA	*dapl1-a*	0.947	0.953	0.946	1.144	0.588	0.684	0.251	0.404		
sp|A3KMU5|DAL1B_XENLA	*dapl1-b*	1.093	1.008	0.991	1.230	0.595	0.690	0.251	0.405		
sp|A8E604|RS3AB_XENLA	*rps3a-b*	0.256	0.560	0.440	0.313	0.413	0.438	0.483	0.374		
sp|P00407|COX2_XENLA	*mt-co2*	-0.305	-0.186	0.062	-0.169	0.658	0.741	0.736	0.496		
sp|P02350|RS31_XENLA	*rps3-a*	0.585	0.802	0.659	0.555	0.605	0.395	0.834	0.571		
sp|P02362|RS7_XENLA	*rps7*	0.060	0.462	0.588	0.866	0.381	0.468	0.684	0.584		
sp|P02377|RS24_XENLA	*rps24*	0.101	0.186	0.667	0.528	0.245	0.561	0.581	0.457		
sp|P03931|ATP8_XENLA	*mt-atp8*	0.921	0.574	0.435	0.912	-1.227	-0.723	0.216	-1.208		
sp|P04751|ACTC_XENLA	***actc1***	**1.728**	**1.323**	**1.413**	**1.573**	**1.403**	**1.252**	**1.584**	**1.293**	**1.77**	**0.020**
sp|P08429|RL4A_XENLA	*rpl4-a*	0.987	0.758	0.546	0.662	0.463	0.320	0.343	0.429		
sp|P09897|RL18A_XENLA	*rpl18-a*	0.667	0.471	0.348	0.440	0.405	0.386	0.398	0.256		
sp|P0DP35|CAM2B_XENLA	*cam2b*	-0.289	-0.869	-0.662	-0.154	0.625	0.737	0.549	0.697		
sp|P15107|SOD1B_XENLA	*sod1-b*	-0.167	-0.287	-0.348	-0.195	0.476	0.506	0.885	0.463		
sp|P17507|EF1A2_XENLA	*eef1ao*	0.809	0.681	0.693	0.874	0.943	0.808	0.944	0.969		

sp|P17508|EF1A3_XENLA	*ef1a3*	0.809	0.681	0.692	0.874	0.955	0.822	0.931	0.964		
sp|P18248|PCNA_XENLA	*pcna*	-1.284	-1.865	-0.958	-1.149	-0.385	-0.546	-1.146	-0.915		
sp|P18709|VITA2_XENLA	*vtga2*	0.767	0.494	0.807	0.431	2.918	3.045	2.854	2.744		
sp|P18758|TYB4_XENLA	*tmsb4*	1.136	0.936	0.884	1.133	0.673	0.360	0.632	0.523		
sp|P19011|VITB2_XENLA	*vtgb2*	-0.346	-0.228	0.127	-0.211	2.580	2.808	2.550	2.508		
sp|P20342|RS15_XENLA	*rps15*	0.524	0.899	0.881	0.660	0.332	0.584	0.652	0.574		
sp|P29309|1433_XENLA	*1433*	-0.235	-0.280	-1.135	-0.194	-0.493	-0.613	-1.323	-0.651		
sp|P29693|EF1D_XENLA	*eef1d*	0.071	0.203	-0.141	0.115	0.033	-0.052	0.376	0.046		
sp|P30151|EF1B_XENLA	*eef1b*	0.420	0.176	0.275	0.614	0.257	0.195	0.006	0.222		
sp|P39017|RS6_XENLA	*rps6*	0.838	0.448	0.399	0.423	0.005	-0.087	0.092	-0.107		
sp|P41116|RL8_XENLA	*rpl8*	2.033	0.909	0.910	1.110	0.627	0.675	1.045	0.579		
sp|P45441|YBX2B_XENLA	*ybx2-b*	0.429	0.029	-0.120	0.429	-0.642	-0.406	-0.140	-0.073		
sp|P45593|COF1B_XENLA	*cfl1-b*	-0.258	-0.048	0.454	0.156	-0.053	0.070	0.792	-0.591		
sp|P45695|COF1A_XENLA	*cfl1-a*	0.348	0.297	0.587	0.321	-0.053	0.070	0.792	-0.534		
sp|P47830|RL27A_XENLA	*rpl27a*	1.086	0.801	0.818	0.808	1.150	1.779	1.429	1.427		
sp|P49393|RS13_XENLA	*rps13*	0.765	0.739	0.654	0.987	0.603	0.332	0.479	0.444		
sp|P49401|RS4_XENLA	*rps4*	0.672	0.615	0.643	0.665	0.523	0.437	0.340	0.430		
sp|P49739|MCM3M_XENLA	*mcm3m*	-2.486	-2.367	-2.859	-2.351	-1.646	-1.543	-0.743	-1.409		
sp|P50143|TCPG_XENLA	*cct3*	-0.620	-0.479	-0.568	-1.244	-0.162	-0.311	-0.240	-0.046		
sp|P50886|RL22_XENLA	***rpl22***	**0.627**	**0.470**	**0.596**	**0.158**	**0.344**	**0.337**	**0.587**	**0.305**	**1.65**	**0.254**
sp|P51893|SAHHA_XENLA	*ahcy-a*	-0.505	-0.163	-0.052	0.057	0.451	0.101	0.244	0.299		
sp|P52170|IMA5_XENLA	*kpna1*	0.171	-0.583	-0.379	-0.062	0.710	0.031	-0.448	0.660		
sp|P52301|RAN_XENLA	*ran*	0.465	0.469	0.663	0.433	0.543	0.330	0.026	0.528		
sp|P55861|MCM2_XENLA	*mcm2*	-0.790	-2.070	-1.862	-1.354	-1.720	-1.580	-1.351	-2.074		
sp|P70010|NDKA1_XENLA	*ndka1*	0.856	1.036	1.016	1.008	1.075	1.010	1.421	1.077		
sp|P79928|P2RY4_XENLA	*p2ry4*	-1.523	-0.986	-0.201	-0.689	-0.547	-2.323	-1.385	-2.809		
sp|P84233|H32_XENLA	*h32*	-0.161	-0.183	-0.560	-0.615	-0.299	-2.250	-1.311	-2.735		
sp|Q07254|RS10_XENLA	*rps10*	0.750	0.415	0.254	0.571	0.268	0.130	0.672	0.502		
sp|Q4U0Y4-2|NPL1A_XENLA	*nap1l1-a*	-1.413	-0.169	-0.272	-0.579	-0.031	-0.384	0.275	0.029		
sp|Q66KU4|RL36_XENLA	*rpl36*	1.053	0.298	0.528	0.762	0.554	0.729	0.067	0.570		
sp|Q6AZJ9|RS21_XENLA	*rps21*	0.985	0.642	0.296	0.929	-0.493	-0.240	0.000	-0.571		
sp|Q6GM74|H2AV_XENLA	*h2afv*	0.549	0.036	-0.102	0.468	-1.733	-1.228	-0.290	-0.397		
sp|Q6GN02|THILB_XENLA	*acat1-b*	-0.255	-0.307	-0.383	0.015	-0.030	-1.754	-0.816	-1.498		

sp|Q6NUH0|RL31_XENLA	*rpl31*	0.579	0.190	0.455	0.550	-0.056	0.000	0.486	0.086		
sp|Q6PA58|SDHAA_XENLA	*sdha-a*	-1.997	-2.577	-1.671	-1.862	-1.242	-1.645	-1.858	-1.159		
sp|Q6PAB3|MDHC_XENLA	*mdhc*	-1.512	-0.450	-0.805	-0.678	0.520	0.438	0.758	0.624		
sp|Q6PAY8|HSDL2_XENLA	*hsdl2*	0.152	-0.369	-1.315	-0.307	-0.412	-0.399	-1.503	-0.885		
sp|Q76BK2|MIF_XENLA	*mif*	-0.381	-0.683	0.000	0.317	-0.130	-0.189	-0.941	-0.671		
sp|Q7ZYF2|TCTP_XENLA	*tpt1*	-1.149	-0.166	-0.567	-0.314	-0.109	-1.949	-1.010	-2.434		
sp|Q7ZYS1|RL19_XENLA	*rpl19*	1.295	0.467	0.563	1.030	0.493	0.181	0.215	0.469		
sp|Q7ZYS8|RL10A_XENLA	*rpl10a*	0.587	0.764	0.654	0.789	0.691	0.590	0.218	0.468		
sp|Q801S3|RS3AA_XENLA	*rps3a-a*	0.666	0.809	0.592	0.647	0.561	0.549	0.577	0.501		
sp|Q91375|EF1GB_XENLA	*eef1g-b*	0.519	0.122	0.000	0.670	-0.334	-0.342	-0.052	-0.303		
sp|Q91896|1433Z_XENLA	*ywhaz*	-1.228	-0.412	-0.902	-0.250	-0.361	-0.584	-0.116	-0.427		
sp|Q9PVQ1|PSA7B_XENLA	*psma7-b*	-0.605	-0.271	-0.395	-0.520	-0.424	-0.165	-1.215	-0.266	**0.61**	**0.209**
TC417616	*hspa9*	-0.318	-0.468	-0.585	-0.540	-1.203	-1.092	-1.236	-1.146		
TC420595	*tubb4b*	0.344	0.872	0.827	0.785	1.007	0.859	1.148	0.975		
TC421129	*ckb*	-0.095	-0.007	0.260	0.128	0.539	0.383	0.846	0.587		
TC424806	*btf3*	-0.489	0.328	0.210	-0.095	0.048	0.231	0.097	0.265		
TC424927	*psma5*	-0.165	-0.198	-0.536	-0.523	-0.598	-0.465	-1.219	-0.635		
TC426340	*rpl9*	0.370	0.056	0.126	0.338	0.198	0.314	0.472	0.314		
TC427469	*cct6a*	-1.331	-1.212	-0.476	-1.196	0.131	-0.062	-0.037	0.125		
TC427650	*rps19*	1.502	1.036	0.914	1.052	0.560	0.430	0.812	0.606		
TC427797	*sdhb*	-0.319	-0.621	-0.806	-0.457	-1.049	-0.637	-1.153	-0.890		
TC433567	*rps9*	0.011	0.464	0.297	-0.467	0.221	0.356	0.163	0.276		
TC435033	*atp5b*	0.656	0.982	0.757	-0.476	0.666	0.709	1.302	0.717		
TC439030	*rps9*	-0.574	0.314	0.065	-1.138	0.081	0.158	0.034	0.149		
TC440338	*rpl27*	0.970	1.043	0.945	0.329	0.868	0.634	0.688	0.724		
TC440824	*hmgb2*	0.185	-0.096	-0.228	0.424	-0.316	-0.080	-0.416	-0.740		
TC441923	*cox6a1*	0.604	0.621	0.419	0.954	0.320	0.102	0.044	0.418		
TC443102	*psma2*	-1.013	-0.438	-0.443	-0.878	0.170	-0.118	-0.176	-0.311		
TC443725	*rpl6*	1.154	0.769	0.656	0.784	0.291	0.176	0.339	0.298		
TC444502	*rps13*	0.679	0.618	0.527	0.956	0.516	0.273	0.479	0.326		
TC446027	*bola2*	0.253	-0.075	-0.701	-0.030	-0.230	-1.827	-0.889	-2.313		
TC446485	*rpl36a*	0.864	0.133	0.098	0.441	-0.226	-0.650	-0.410	-0.184		
TC446731	*rpl14*	1.011	0.656	0.541	0.432	1.385	1.098	0.525	0.508		

TC446993	*vcp*	-1.541	-0.413	-0.378	-0.565	-0.087	-0.124	-0.057	-0.078		
TC447912	*stmn1*	0.619	0.363	0.076	0.354	-0.115	0.166	-0.341	-0.290		
TC449759	*ckb*	0.728	0.565	0.631	0.668	0.562	0.459	0.781	0.619		
TC449989	*atp5o*	-1.707	-0.140	-0.345	-0.872	0.098	0.008	0.209	-0.233		
TC450779	*rps26*	0.234	0.071	0.059	0.651	0.549	0.314	0.784	0.573		
TC451297	*uqcrc1*	-0.599	-1.414	-0.996	-0.460	-1.491	-2.332	-1.394	-1.819		
TC453026	*hist1h2bj*	0.558	0.314	0.315	0.576	0.300	-0.118	0.124	0.226		
TC455791	*rps10*	0.822	0.394	0.366	0.658	0.302	0.144	0.672	0.503		
TC456558	*anxa5*	-1.932	-1.813	-1.120	-1.796	-0.557	-0.758	-1.793	-0.443		
TC456737	*rpl35*	0.859	0.901	0.780	0.224	0.450	0.390	0.539	0.416		
TC457286	*vtga2*	1.002	0.944	1.217	0.366	1.395	1.579	1.331	1.292		
TC458492	*acaa2*	-0.655	-0.028	-0.071	0.000	-0.954	-0.905	0.255	-1.115		
TC460552	*hspe1*	1.563	0.832	0.426	1.085	0.211	0.224	0.394	0.199		
TC461678	*aldoa*	0.821	0.742	0.639	0.617	1.143	0.893	0.912	0.951		
TC462321	*vtga2*	-1.454	0.341	0.290	-0.620	2.967	3.126	2.947	2.854		
TC462743	*vtgb1*	-0.539	-0.329	-0.033	-0.890	2.830	3.004	2.815	2.714		
TC463468	*vtga1*	-1.195	0.245	0.332	-0.361	2.650	2.918	2.623	2.464		
TC464477	*ptpn11*	0.002	0.051	-0.144	-0.371	-2.509	-2.005	-1.066	-2.491		
TC464539	*vtgb1*	-0.117	0.047	0.216	-0.007	2.676	2.753	2.503	2.400		
TC464706	*fau*	1.120	0.997	0.942	1.379	-0.409	-0.604	0.335	-1.089		
TC466790	*adrm1*	0.544	-0.002	-0.229	0.441	-1.860	-1.355	-0.417	-1.841		
TC467271	*slc25a5*	-1.030	-0.425	-0.991	0.192	0.194	0.509	0.196	0.175		
TC467404	*rps12*	0.354	0.437	0.242	0.285	0.340	0.327	0.790	0.141		
TC469090	*vtga2*	1.047	-0.732	-0.961	1.198	2.055	2.109	1.929	1.777		
zeinaSSns_comp323944_c0_seq2	***pcmt1***	**-0.944**	**-0.532**	**-0.618**	**0.162**	**-0.277**	**0.101**	**-0.806**	**-0.385**	**0.34**	**0.013**
zeinaSSns_comp358456_c3_seq1	*unknown*	0.689	0.527	0.483	0.695	0.408	0.497	0.573	0.263		
zeinaSSns_comp359224_c0_seq1	*atp5i*	0.163	0.128	-0.147	0.318	-1.840	-0.331	-0.397	-0.610		
zeinaSSns_comp369857_c0_seq1	*rpl13*	1.591	1.000	0.820	1.080	0.746	0.528	1.136	0.707		
zeinaSSns_comp370410_c0_seq1	*dbi*	0.600	0.638	0.531	0.905	0.166	0.504	0.040	0.214		
zeinaSSns_comp370555_c0_seq1	***rpl21***	**0.317**	**0.333**	**0.238**	**-0.039**	**0.318**	**0.207**	**0.680**	**0.234**	**1.38**	**0.009**
zeinaSSns_comp374146_c0_seq3	*nasp*	-0.813	-0.411	-0.756	-0.259	-0.124	-0.405	-0.413	-0.141		
zeinaSSns_comp384836_c0_seq2	*gyg1*	-1.250	-1.097	-0.559	-0.415	-1.152	-1.129	-1.111	-1.242		
zeinaSSns_comp385614_c2_seq7	*eef1d*	0.491	0.320	0.113	0.513	0.181	0.022	0.312	0.075		

zeinaSSns_comp386237_c1_seq2	*h2ac21*	0.498	0.339	0.278	0.944	-0.494	-0.472	-0.232	-0.187		
zeinaSSns_comp386551_c0_seq4	*wdr1*	-1.459	-2.040	-1.832	-1.324	-0.599	-0.387	-0.622	-0.721		
zeinaSSns_comp386737_c5_seq20	*rps25*	0.769	0.925	0.652	0.962	0.192	0.764	0.867	0.558		
zeinaSSns_comp387272_c0_seq1	***eno1***	**1.141**	**1.100**	**1.236**	**0.969**	**1.258**	**1.063**	**1.407**	**1.119**	**1.60**	**0.166**
zeinaSSns_comp387533_c0_seq4	*vat1*	-1.017	-1.597	-0.691	-0.881	-0.927	-0.274	-0.878	-0.804		
zeinaSSns_comp389333_c5_seq4	*rpl37a*	0.209	0.091	0.222	-0.355	0.148	0.207	0.577	0.284		
zeinaSSns_comp389630_c0_seq1	*kpnb1*	-0.941	-1.064	-1.022	-1.533	-0.232	-0.543	-0.831	-0.272		

Molecular features (*m/z*, t$_m$ in min)	**Metabolite Name**	**E1**		**E2**		**E3**		**E4**		**Mean FC**	***p*-value**
		L-D1$_1$	**L-V1$_1$**	**L-D1$_2$**	**L-V1$_2$**	**L-D1$_3$**	**L-V1$_3$**	**L-D1$_4$**	**L-V1$_4$**		
89.1070, 9.3	Putrescine	-0.350	-0.831	-1.060	-0.353	-0.462	-0.580	-1.306	-0.828		
90.0548, 20.2	Alanine	-0.299	-0.645	-0.218	-0.382	-0.130	0.137	-0.293	-0.568		
90.0552, 15.3	Sarcosine	-1.222	-1.589	-1.668	-1.851	-1.317	-1.405	-1.652	-1.845		
104.0707, 17.3	y-aminobutyrate	-0.332	-0.829	-0.180	-0.202	-0.280	-0.243	-0.372	-0.439		
104.0707, 17.7		1.791	1.626	1.581	1.748	1.697	1.586	1.248	1.651		
104.1066, 14.1	Choline	0.179	-0.432	0.066	0.109	-0.074	-0.071	0.872	0.765		
106.0493, 24.7	Serine	0.198	0.168	0.266	0.307	0.292	0.311	0.514	0.186		
110.0696, 16.3		-0.430	-0.361	-0.302	-0.371	-0.323	-0.222	-0.158	-0.403		
114.0649, 16.1		**-0.505**	**-0.621**	**-0.203**	**-0.894**	**-0.720**	**-1.330**	**-0.102**	**-1.053**	**4.81**	**0.043**
116.0703, 28.6	Proline	0.504	0.387	0.334	0.377	0.519	0.612	0.725	0.469		
118.0856, 30.8	**Glycine betaine**	**0.064**	**-0.234**	**0.042**	**-0.058**	**0.161**	**0.107**	**0.302**	**0.000**	**1.59**	**0.062**
118.0853, 24.6	Valine	-0.280	0.259	0.090	0.072	0.068	0.159	0.364	0.005		
120.0650, 26.2	Threonine	0.093	0.126	0.172	0.063	0.211	0.261	0.285	0.201		
122.0813, 18.6		2.645	2.505	2.437	3.310	2.543	2.478	2.131	2.463		
123.0447, 11.0		-0.673	-0.507	-0.548	-0.236	-0.394	-0.435	-0.733	-0.626		
126.1019, 9.8	Methylhistamine	0.045	-1.931	0.054	-0.056	0.000	0.131	0.227	0.257		
130.0503, 28.3		0.806	0.385	0.920	1.032	0.855	0.919	0.880	0.796		
130.0503, 28.8		**1.336**	**0.838**	**1.176**	**0.874**	**1.234**	**1.283**	**1.121**	**1.023**	**1.83**	**0.173**
130.1583, 20.9		0.081	0.299	-0.319	-0.217	-0.073	-0.377	-0.241	-0.382		
132.0652, 32.6	**Hydroxyproline**	**0.150**	**0.112**	**0.157**	**0.137**	**0.250**	**0.130**	**0.157**	**0.087**	**1.16**	**0.066**
132.0767, 20.2	Creatine	1.157	0.928	1.263	1.102	1.297	1.613	1.172	0.935		

132.1014, 24.9	Isoleucine	-0.309	0.341	-0.091	0.149	-0.020	-0.018	0.226	-0.082		
132.1014, 25.1	Leucine	0.000	0.432	0.180	0.233	0.213	0.221	0.328	0.076		
133.0341, 11.1		1.272	1.667	0.928	1.160	1.100	1.168	0.978	0.845		
133.0612, 26.2	**Asparagine**	**-0.076**	**-0.259**	**-0.045**	**-0.208**	**-0.044**	**0.013**	**0.011**	**-0.342**	**1.53**	**0.152**
133.0983, 15.1	**Ornithine**	**-0.342**	**0.492**	**-0.142**	**0.139**	**-0.039**	**-0.004**	**0.275**	**0.112**	**0.76**	**0.336**
134.0445, 30.2	Aspartate	1.322	1.360	1.332	1.289	1.435	1.493	1.414	1.360		
134.1169, 18.4		0.194	-0.459	-0.024	0.131	0.263	0.000	-0.022	0.067		
136.0766, 29.7		-0.478	-1.311	-0.516	-0.553	-0.403	-0.586	-0.397	-0.424		
137.0293, 11.3		1.254	1.583	1.023	1.169	1.051	1.130	0.934	0.801		
137.0454, 28.6	Hypoxanthine	1.380	0.832	1.391	1.235	1.378	1.385	1.065	1.337		
139.0505, 17.5	Urocanate, -trans	-1.744	0.333	-0.376	-0.182	-0.654	-0.245	0.313	-0.291		
142.1222, 18.3		0.365	0.000	0.752	0.453	0.498	0.813	0.443	0.313		
146.1173, 16.9	Acetylcholine	0.309	-0.006	0.292	0.253	0.350	0.361	0.250	0.159		
147.013, 11.30		0.287	0.676	0.036	0.197	0.155	0.212	0.083	-0.087		
147.0766, 26.9	Glutamine	0.956	0.490	1.086	1.020	0.998	1.066	1.039	0.959		
147.1129, 15.3	Lysine	0.699	0.611	0.626	0.626	0.742	0.770	0.544	0.509		
148.0604, 19.9	Methylaspartate	0.301	0.269	0.328	0.266	1.491	0.284	0.155	0.092		
148.0609, 27.3	Glutamate	1.458	1.376	1.747	1.716	1.791	1.834	1.680	1.578		
148.1139, 15.5		-0.404	-0.585	-0.492	-0.441	-0.346	-0.336	-0.562	-0.593		
150.0548, 20.6		0.190	-1.228	0.149	0.123	0.306	-0.759	-0.287	0.385		
150.0594, 26.8	**Methionine**	**0.131**	**-0.664**	**0.135**	**-0.183**	**0.160**	**0.049**	**0.161**	**0.199**	**2.63**	**0.201**
150.1134, 17.1	**Trolamine**	**-0.651**	**1.656**	**-1.089**	**0.449**	**-0.559**	**-0.666**	**-0.468**	**-0.645**	**0.70**	**0.244**
152.046, 11.3		0.506	0.756	0.243	0.408	0.273	0.382	-0.094	0.056		
152.0569, 17.5	Guanine	1.774	0.788	1.767	1.632	1.795	1.613	1.469	1.899		
152.1181, 10.6		-0.520	-0.833	0.112	-0.287	-0.309	0.402	-0.170	-0.279		
156.0768, 16.0	Histidine	0.642	0.645	0.694	0.664	0.699	0.802	0.796	0.591		
157.1704, 12.2		-0.109	1.900	-0.098	0.067	0.806	0.050	-0.339	0.039		
158.1537, 22		-1.182	0.530	-0.783	-0.825	-0.384	-0.673	-1.014	-0.914		
162.0758, 27.0	2-Aminoadipate	-0.981	-0.321	-0.887	-1.292	-1.178	-1.145	-1.003	-0.953		
162.1128, 17.9	Carnitine	1.340	1.312	1.307	1.214	1.378	1.383	1.248	1.143		
164.0459, 11.3		-0.473	-1.007	-0.414	-0.164	-0.322	-0.379	-0.415	-0.602		
165.0557, 29.8		0.015	-0.455	-0.037	-0.060	0.034	-0.046	0.003	-0.030		
166.0547, 29.7		-0.894	-0.931	-0.994	-1.089	-1.012	-1.040	-1.018	-1.063		

166.0866, 35.3		0.651	0.524	0.675	0.666	0.729	0.723	0.711	0.624		
168.0654, 18.4	Pyridoxal (Vitamin B6)	0.271	-0.394	0.217	0.217	0.372	-0.446	-0.134	0.418		
172.0977, 21.4		-1.712	-0.098	-1.734	-2.141	-1.344	-1.756	-1.594	-2.192		
175.1185, 15.8	Arginine	0.635	0.847	0.637	0.680	0.790	0.783	0.541	0.573		
176.1025, 27.3	Citrulline	-0.467	-0.123	-0.278	-0.257	-0.311	-0.349	-0.010	-0.363		
182.0806, 28.2	Tyrosine	0.437	0.000	0.361	0.367	0.429	0.380	0.370	0.360		
182.1885, 21.7		-0.594	-0.930	-0.441	-1.417	-0.756	-1.007	-1.267	-1.080	**3.52**	**0.247**
184.1092, 17.8		**-1.684**	**0.035**	**-1.668**	**-1.760**	**-2.324**	**-1.287**	**-1.761**	**-1.573**	**0.50**	**0.182**
185.1639, 20.2		-0.639	1.455	-0.774	-1.144	0.000	-1.634	-1.559	-0.595		
187.1424, 19.5		-0.385	-0.185	-0.445	0.000	-0.258	-0.551	-1.099	-0.262		
188.0696, 33.4		**-0.045**	**-0.446**	**0.000**	**-0.099**	**0.052**	**-0.060**	**0.000**	**0.000**	**1.52**	**0.174**
189.1592, 16.0	**N6, N6, N6-trimethyl-lysine**	**0.078**	**-0.078**	**0.077**	**-0.022**	**0.077**	**0.135**	**0.135**	**0.019**	**1.22**	**0.197**
190.0551, 17.7		-0.911	-0.887	-0.992	-0.927	-0.510	-1.102	-1.318	-1.100		
190.1178, 27.3	**Homocitrulline**	**-1.220**	**-1.962**	**-1.222**	**-1.510**	**-1.254**	**-1.364**	**-1.238**	**-1.179**	**2.40**	**0.215**
191.1026, 11.0		0.611	0.470	0.446	0.701	0.764	0.566	0.306	0.441		
194.1553, 21.6		-0.151	-1.311	-0.350	-0.336	-0.317	-0.434	-0.265	-0.015		
202.1804, 22.4		-1.397	0.709	-1.729	-1.710	-0.509	-1.792	-0.383	-1.559		
204.1227, 19.3	Acetylcarnitine	0.382	0.155	0.236	0.189	0.247	0.349	0.171	0.172		
205.1181, 29.0	Ser-Val	-1.089	-1.154	-1.080	-1.334	-1.026	-0.964	-0.849	-0.823		
208.1804, 12.3		-0.702	-0.228	-1.215	-1.010	-1.112	-1.015	-1.461	-1.201		
212.2002, 21.3		-0.147	-0.372	-0.074	-0.268	-0.045	-0.402	-0.369	-0.008		
230.2469, 23.6		0.128	-0.133	0.248	-0.698	-0.112	0.378	-0.929	-0.562		
255.0971, 7.7		-0.503	-1.169	-0.887	0.692	-0.573	-0.599	-0.772	-0.686		
255.0971, 22.7		-0.616	-0.889	-0.592	-0.706	-0.647	-0.683	-0.561	-0.392		
262.1513, 14.9	Ser-Arg	-1.408	-1.195	-1.407	-1.580	-1.312	-1.170	-1.751	-1.096		
291.1295, 22.7	Argininosuccinate	-0.435	-0.824	-0.750	-0.798	-0.637	-0.588	-1.046	-1.042		
307.1218, 25.2		-0.569	-0.724	-0.600	-0.723	-0.629	-0.726	-0.684	-0.509		
308.0909, 34.4	Glutathione, reduced	1.205	1.163	1.181	1.074	1.202	1.285	1.034	0.922		
359.2107, 10.0		-0.738	0.039	-0.899	-1.063	-0.782	-0.590	-1.540	-1.337		

Bibliography

1. Rinschen, M. M.; Ivanisevic, J.; Giera, M.; Siuzdak, G., Identification of bioactive metabolites using activity metabolomics. *Nat. Rev. Mol. Cell Biol.* **2019,** *20* (6), 353-367.
2. Markley, J. L.; Brüschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S., The future of NMR-based metabolomics. *Curr. Opin. Biotechnol.* **2017,** *43*, 34-40.
3. Liu, H. L.; Hsu, J. P., Recent developments in structural proteomics for protein structure determination. *Proteomics* **2005,** *5* (8), 2056-2068.
4. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.
5. Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* **2007,** *26* (1), 51-78.
6. Onjiko, R. M.; Moody, S. A.; Nemes, P., Single-cell mass spectrometry reveals small molecules that affect cell fates in the 16-cell embryo. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112* (21), 6545-6550.
7. Vastag, L.; Jorgensen, P.; Peshkin, L.; Wei, R.; Rabinowitz, J. D.; Kirschner, M. W., Remodeling of the metabolome during early frog development. *PLoS ONE* **2011,** *6* (2), e16881.
8. Choi, S. B.; Zamarbide, M.; Manzini, M. C.; Nemes, P., Tapered-tip capillary electrophoresis nano-electrospray ionization mass spectrometry for ultrasensitive proteomics: the mouse cortex. *J. Am. Soc. Mass Spectrom.* **2016,** *28* (4), 597-607.
9. de Souza, L. P.; Alseekh, S.; Scossa, F.; Fernie, A. R., Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. *Nat. Methods* **2021,** *18* (7), 733-746.
10. Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ. Sci. Technol.* **2014,** *48* (4), 2097-2098.

11. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN: a metabolite mass spectral database. *Ther. Drug Monit.* **2005,** *27* (6), 747-751.

12. Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; McLean, J. A., Untargeted metabolomics strategies-challenges and emerging directions. *J. Am. Soc. Mass Spectrom.* **2016,** *27* (12), 1897-1905.

13. Forsberg, E. M.; Huan, T.; Rinehart, D.; Benton, H. P.; Warth, B.; Hilmers, B.; Siuzdak, G., Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS online. *Nat. Protoc.* **2018,** *13* (4), 633-651.

14. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010,** *11* (1), 1-11.

15. Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **2011,** *10* (4), 1794-1805.

16. Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008,** *26* (12), 1367-1372.

17. Orsburn, B. C., Proteome Discoverer-a community enhanced data processing suite for protein informatics. *Proteomes* **2021,** *9* (1), 1-13.

18. Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H., Ion mobility-mass spectrometry. *J. Mass Spectrom.* **2008,** *43* (1), 1-22.

19. Fouque, K. J. D.; Fernandez-Lima, F., Recent advances in biological separations using trapped ion mobility spectrometry - mass spectrometry. *Trac-Trends Anal. Chem.* **2019,** *116*, 308-315.

20. DeLaney, K.; Jia, D. S.; Iyer, L.; Yu, Z.; Choi, S. B.; Marvar, P. J.; Nemes, P., Microanalysis of brain angiotensin peptides using ultrasensitive capillary electrophoresis trapped ion mobility mass spectrometry. *Anal. Chem.* **2022,** *94* (25), 9018-9025.

21. Nys, G.; Cobraiville, G.; Fillet, M., Multidimensional performance assessment of micro pillar array column chromatography combined to ion mobility-mass spectrometry for proteome research. *Anal. Chim. Acta* **2019,** *1086*, 1-13.

22. Baxi, A. B.; Lombard-Banek, C.; Moody, S. A.; Nemes, P., Proteomic characterization of the neural ectoderm fated cell clones in the *Xenopus laevis* embryo by high-resolution mass spectrometry. *ACS Chem. Neurosci.* **2018,** *9* (8), 2064–2073.
23. Mertins, P.; Tang, L. C.; Krug, K.; Clark, D. J.; Gritsenko, M. A.; Chen, L. J.; Clauser, K. R.; Clauss, T. R.; Shah, P.; Gillette, M. A.; Petyuk, V. A.; Thomas, S. N.; Mani, D. R.; Mundt, F.; Moore, R. J.; Hui, Y. W.; Zhao, R.; Schnaubelt, M.; Keshishian, H.; Monroe, M. E.; Zhang, Z.; Udeshi, N. D.; Mani, D.; Davies, S. R.; Townsend, R. R.; Chan, D. W.; Smith, R. D.; Zhang, H.; Liu, T.; Carr, S. A., Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat. Protoc.* **2018,** *13* (7), 1632-1661.
24. Keshishian, H.; Burgess, M. W.; Gillette, M. A.; Mertins, P.; Clauser, K. R.; Mani, D. R.; Kuhn, E. W.; Farrell, L. A.; Gerszten, R. E.; Carr, S. A., Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol. Cell. Proteomics* **2015,** *14* (9), 2375-2393.
25. Demirev, P. A.; Fenselau, C., Mass spectrometry for rapid characterization of microorganisms. *Annu. Rev. Anal. Chem.* **2008,** *1*, 71-93.
26. Lombard-Banek, C.; Li, J.; Portero, E. P.; Onjiko, R. M.; Singer, C. D.; Plotnick, D. O.; Al Shabeeb, R. Q.; Nemes, P., In vivo subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*). *Angew. Chem. Int. Edit.* **2021,** *60* (23), 12852-12858.
27. Liu, J. X.; Aerts, J. T.; Rubakhin, S. S.; Zhang, X. X.; Sweedler, J. V., Analysis of endogenous nucleotides by single cell capillary electrophoresis-mass spectrometry. *Analyst* **2014,** *139* (22), 5835-5842.
28. Zhang, L. W.; Khattar, N.; Kemenes, I.; Kemenes, G.; Zrinyi, Z.; Pirger, Z.; Vertes, A., Subcellular peptide localization in single identified neurons by capillary microsampling mass spectrometry. *Sci Rep* **2018,** *8* (1), 1-10.
29. Zhang, L.; Vertes, A., Single-cell mass spectrometry approaches to explore cellular heterogeneity. *Angew. Chem. Int. Edit.* **2018,** *57* (17), 4466-4477.
30. Lombard-Banek, C.; Moody, S. A.; Nemes, P., Single-cell mass spectrometry for discovery proteomics: quantifying translational cell heterogeneity in the 16-cell frog (*Xenopus*) embryo. *Angew. Chem. Int. Edit.* **2016,** *128* (7), 2500-2504.

31. Jang, C.; Chen, L.; Rabinowitz, J. D., Metabolomics and Isotope Tracing. *Cell* **2018,** *173* (4), 822-837.

32. Baxi, A. B.; Pade, L. R.; Nemes, P., Cell-lineage guided mass spectrometry proteomics in the developing (frog) embryo. *J. Vis. Exp.* **2022,** (182), 22.

33. Gatto, L.; Aebersold, R.; Cox, J.; Demichev, V.; Derks, J.; Emmott, E.; Franks, A. M.; Ivanov, A. R.; Kelly, R. T.; Khoury, L.; Leduc, A.; MacCoss, M. J.; Nemes, P.; Perlman, D. H.; Petelski, A. A.; Rose, C. M.; Schoof, E. M.; Van Eyk, J.; Vanderaa, C.; Yates, J. R., III; Slavov, N. Initial recommendations for performing, benchmarking, and reporting single-cell proteomics experiments **2022**, p. 10815. (accessed July 01, 2022).

34. Sugimoto, M.; Wong, D. T.; Hirayama, A.; Soga, T.; Tomita, M., Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics* **2010,** *6* (1), 78-95.

35. Mellick, G. D.; Silburn, P. A.; Sutherland, G. T.; Siebert, G. A., Exploiting the potential of molecular profiling in Parkinson's disease: current practice and future probabilities. *Expert Rev. Mol. Diagn.* **2010,** *10* (8), 1035-1050.

36. Lei, Z.; Huhman, D. V.; Sumner, L. W., Mass spectrometry strategies in metabolomics. *J. Biol. Chem.* **2011,** *286* (29), 25435-25442.

37. Patti, G. J.; Yanes, O.; Siuzdak, G., Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **2012,** *13* (4), 263-269.

38. Zhu, Z. J.; Schultz, A. W.; Wang, J.; Johnson, C. H.; Yannone, S. M.; Patti, G. J.; Siuzdak, G., Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. *Nat. Protoc.* **2013,** *8* (3), 451-460.

39. Sandra, K.; Vandenbussche, J.; Vandenheede, I.; Claerebout, B.; Op de Beeck, J.; Jacobs, P.; De Malsche, W.; Desmet, G.; Sandra, P., Peptide mapping of monoclonal antibodies and antibody-drug conjugates using micro-pillar array columns combined with mass spectrometry. *LC GC Eur.* **2018,** *31* (3), 155-166.

40. Rozing, G., Micropillar array columns for advancing nanoflow HPLC. *Microchem. J.* **2021,** *170*, 106629.

41. Nemes, P.; Rubakhin, S. S.; Aerts, J. T.; Sweedler, J. V., Qualitative and quantitative metabolomic investigation of single neurons by capillary electrophoresis electrospray ionization mass spectrometry. *Nat. Protoc.* **2013,** *8* (4), 783-799.

42. Nemes, P.; Knolhoff, A. M.; Rubakhin, S. S.; Sweedler, J. V., Metabolic differentiation of neuronal phenotypes by single-cell capillary electrophoresis-electrospray ionization-mass spectrometry. *Anal. Chem.* **2011,** *83* (17), 6810-6817.

43. Onjiko, R. M.; Portero, E. P.; Moody, S. A.; Nemes, P., *In situ* microprobe single-cell capillary electrophoresis mass spectrometry: metabolic reorganization in single differentiating cells in the live vertebrate (*Xenopus laevis*) embryo. *Anal. Chem.* **2017,** *89* (13), 7069-7076.

44. Wang, Y. X.; Yang, F.; Gritsenko, M. A.; Wang, Y. C.; Clauss, T.; Liu, T.; Shen, Y. F.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klemke, R. L.; Camp, D. G.; Smith, R. D., Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* **2011,** *11* (10), 2019-2026.

45. Josic, D.; Kovac, S., Reversed-phase high performance liquid chromatography of proteins. *Current Protocols in Protein Science* **2010,** *61* (1), 8.7.1-8.7.22.

46. Xu, Y., Tutorial: capillary electrophoresis. *The Chemical Educator* **1996,** *1* (2), 1-14.

47. Kennedy, R. T.; Oates, M. D.; Cooper, B. R.; Nickerson, B.; Jorgenson, J. W., Microcolumn separations and the analysis of single cells. *Science* **1989,** *246* (4926), 57-63.

48. Zenobi, R., Single-cell metabolomics: analytical and biological perspectives. *Science* **2013,** *342* (6163), 1243259.

49. Onjiko, R. M.; Portero, E. P.; Moody, S. A.; Nemes, P., Microprobe capillary electrophoresis mass spectrometry for single-cell metabolomics in live frog (*Xenopus laevis*) embryos. *J. Vis. Exp.* **2017,** (130), e56956.

50. Portero, E. P.; Nemes, P., Dual cationic-anionic profiling of metabolites in a single identified cell in a live *Xenopus laevis* embryo by microprobe CE-ESI-MS. *Analyst* **2019,** *144* (3), 892-900.

51. Lombard-Banek, C.; Reddy, S.; Moody, S. A.; Nemes, P., Label-free quantification of proteins in single embryonic cells with neural fate in the cleavage-stage frog (*Xenopus*

laevis) embryo using capillary electrophoresis electrospray ionization high-resolution mass spectrometry (CE-ESI-HRMS). *Mol. Cell. Proteomics* **2016,** *15* (8), 2756-2768.

52. Lombard-Banek, C.; Moody, S. A.; Manzin, M. C.; Nemes, P., Microsampling capillary electrophoresis mass spectrometry enables single-cell proteomics in complex tissues: developing cell clones in live *Xenopus laevis* and zebrafish embryos. *Anal. Chem.* **2019,** *91* (7), 4797-4805.
53. Lombard-Banek, C.; Choi, S. B.; Nemes, P., Single-cell proteomics in complex tissues using microprobe capillary electrophoresis mass spectrometry. In *Methods Enzymol.*, Allbritton, N. L.; Kovarik, M. L., Eds. Academic Press: **2019**; Vol. 628, pp 263-292.
54. Thomas, S.; Hao, L.; Ricke, W. A.; Li, L. J., Biomarker discovery in mass spectrometry-based urinary proteomics. *Proteom. Clin. Appl.* **2016,** *10* (4), 358-370.
55. Wiest, M. M.; Watkins, S. M., Biomarker discovery using high-dimensional lipid analysis. *Curr. Opin. Lipidol.* **2007,** *18* (2), 181-186.
56. Sobsey, C. A.; Ibrahim, S.; Richard, V. R.; Gaspar, V.; Mitsa, G.; Lacasse, V.; Zahedi, R. P.; Batist, G.; Borchers, C. H., Targeted and untargeted proteomics approaches in biomarker development. *Proteomics* **2020,** *20* (9), 1900029.
57. Zhou, F.; Narasimhan, V.; Shboul, M.; Chong, Y. L.; Reversade, B.; Roy, S., Gmnc Is a master regulator of the multiciliated cell differentiation program. *Curr. Biol.* **2015,** *25* (24), 3267-3273.
58. Zhang, L.; Sevinsky, C. J.; Davis, B. M.; Vertes, A., Single-cell mass spectrometry of subpopulations selected by fluorescence microscopy. *Anal. Chem.* **2018,** *90* (7), 4626-4634.
59. Portero, E. P. Development of single-cell mass spectrometry tools to investigate metabolic reorganization during early embryogenesis. Ph.D. Dissertation, University of Maryland, College Park, **2020**.
60. Li, H. R.; Uittenbogaard, M.; Navarro, R.; Ahmed, M.; Gropman, A.; Chiaramello, A.; Hao, L., Integrated proteomic and metabolomic analyses of the mitochondrial neurodegenerative disease MELAS. *Mol. Omics* **2022,** *18* (3), 196-205.
61. Di Girolamo, F.; Lante, I.; Muraca, M.; Putignani, L., The role of mass spectrometry in the "omics" era. *Curr. Org. Chem.* **2013,** *17* (23), 2891-2905.

62. Bock, C.; Farlik, M.; Sheffield, N. C., Multi-omics of single cells: strategies and applications. *Trends Biotechnol.* **2016,** *34* (8), 605-608.

63. Pinu, F. R.; Beale, D. J.; Paten, A. M.; Kouremenos, K.; Swarup, S.; Schirra, H. J.; Wishart, D., Systems biology and multi-omics integration: viewpoints from the metabolomics research community. *Metabolites* **2019,** *9* (4), 76.

64. Ma, A.; McDermaid, A.; Xu, J.; Chang, Y.; Ma, Q., Integrative methods and practical challenges for single-cell multi-omics. *Trends Biotechnol.* **2020,** *38* (9), 1007-1022.

65. Hasin, Y.; Seldin, M.; Lusis, A., Multi-omics approaches to disease. *Genome Biol.* **2017,** *18* (1), 1-15.

66. Lloyd-Price, J.; Arze, C.; Ananthakrishnan, A. N.; Schirmer, M.; Avila-Pacheco, J.; Poon, T. W.; Andrews, E.; Ajami, N. J.; Bonham, K. S.; Brislawn, C. J.; Casero, D.; Courtney, H.; Gonzalez, A.; Graeber, T. G.; Hall, A. B.; Lake, K.; Landers, C. J.; Mallick, H.; Plichta, D. R.; Prasad, M.; Rahnavard, G.; Sauk, J.; Shungin, D.; Vazquez-Baeza, Y.; White, R. A.; Braun, J.; Denson, L. A.; Jansson, J. K.; Knight, R.; Kugathasan, S.; McGovern, D. P. B.; Petrosino, J. F.; Stappenbeck, T. S.; Winter, H. S.; Clish, C. B.; Franzosa, E. A.; Vlamakis, H.; Xavier, R. J.; Huttenhower, C.; Bishai, J.; Bullock, K.; Deik, A.; Dennis, C.; Kaplan, J. L.; Khalili, H.; McIver, L. J.; Moran, C. J.; Nguyen, L.; Pierce, K. A.; Schwager, R.; Sirota-Madi, A.; Stevens, B. W.; Tan, W.; ten Hoeve, J. J.; Weingart, G.; Wilson, R. G.; Yajnik, V.; Investigators, I., Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **2019,** *569* (7758), 655-662.

67. Zhang, W. W.; Li, F.; Nie, L., Integrating multiple 'omics' analysis for microbial biology: application and methodologies. *Microbiology-(UK)* **2010,** *156*, 287-301.

68. Franzosa, E. A.; Hsu, T.; Sirota-Madi, A.; Shafquat, A.; Abu-Ali, G.; Morgan, X. C.; Huttenhower, C., Sequencing and beyond: integrating molecular 'omics' for microbial community profiling. *Nat. Rev. Microbiol.* **2015,** *13* (6), 360-372.

69. Flachsova, M.; Sindelka, R.; Kubista, M., Single blastomere expression profiling of *Xenopus laevis* embryos of 8 to 32-cells reveals developmental asymmetry. *Sci. Rep.* **2013,** *3* (1), 1-6.

70. Quiros, P. M.; Prado, M. A.; Zamboni, N.; D'Amico, D.; Williams, R. W.; Finley, D.; Gygi, S. P.; Auwerx, J., Multi-omics analysis identifies ATF4 as a key regulator of the mitochondrial stress response in mammals. *J. Cell Biol.* **2017,** *216* (7), 2027-2045.
71. Wheeler, G. N.; Brandli, A. W., Simple vertebrate models for chemical genetics and drug discovery screens: lessons from zebrafish and *Xenopus*. *Dev. Dyn.* **2009,** *238* (6), 1287-1308.
72. Moody, S. A., Fates of the blastomeres of the 16-cell stage *Xenopus* embryo. *Dev. Biol.* **1987,** *119* (2), 560-578.
73. Moody, S. A., Fates of the blastomeres of the 32-cell-stage *Xenopus* embryo. *Dev. Biol.* **1987,** *122* (2), 300-319.
74. Peuchen, E. H.; Sun, L.; Dovichi, N. J., Optimization and comparison of bottom-up proteomic sample preparation for early-stage *Xenopus laevis* embryos. *Anal. Bioanal. Chem.* **2016,** *408* (17), 4743-4749.
75. Karpinka, J. B.; Fortriede, J. D.; Burns, K. A.; James-Zorn, C.; Ponferrada, V. G.; Lee, J.; Karimi, K.; Zorn, A. M.; Vize, P. D., Xenbase, the *Xenopu*s model organism database; new virtualized system, data types and genomes. *Nucleic Acids Res.* **2015,** *43* (43), D756-D763.
76. Nemes, P., Mass spectrometry comes of age for subcellular organelles. *Nat. Methods* **2021,** *18* (10), 1157-1158.
77. Dohla, J.; Kuuluvainen, E.; Gebert, N.; Amaral, A.; Englund, J. I.; Gopalakrishnan, S.; Konovalova, S.; Nieminen, A. I.; Salminen, E. S.; Munumer, R. T.; Ahlqvist, K.; Yang, Y.; Bui, H.; Otonkoski, T.; Kakela, R.; Hietakangas, V.; Tyynismaa, H.; Ori, A.; Katajisto, P., Metabolic determination of cell fate through selective inheritance of mitochondria. *Nat. Cell Biol.* **2022,** *24* (2), 148-154.
78. Portero, E. P.; Pade, L. R.; Li, J.; Choi, S. B.; Nemes, P., Single-cell mass spectrometry of metabolites and proteins for systems and functional biology. In *Single Cell 'Omics of Neuronal Cells*, Sweedler, J. V.; Eberwine, J.; Fraser, S. E., Eds. Springer US: New York, NY, **2022**; pp 87-114.
79. Briggs, J. A.; Weinreb, C.; Wagner, D. E.; Megason, S.; Peshkin, L.; Kirschner, M. W.; Klein, A. M., The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **2018,** *360* (6392), eaar5780.

80. Tan, T. Z.; Heong, V.; Ye, J.; Lim, D.; Low, J.; Choolani, M.; Scott, C.; Tan, D. S. P.; Huang, R. Y.-J., Decoding transcriptomic intra-tumour heterogeneity to guide personalised medicine in ovarian cancer. *The Journal of Pathology* **2019,** *247* (3), 305-319.

81. Rohrback, S.; Siddoway, B.; Liu, C. S.; Chun, J., Genomic mosaicism in the developing and adult brain. *Developmental Neurobiology* **2018,** *78* (11), 1026-1048.

82. Cembrowski, M. S., Single-cell transcriptomics as a framework and roadmap for understanding the brain. *J. Neurosci. Methods* **2019,** *326*, 1-7.

83. Phillips, J. W.; Schulmann, A.; Hara, E.; Winnubst, J.; Liu, C.; Valakh, V.; Wang, L.; Shields, B. C.; Korff, W.; Chandrashekar, J.; Lemire, A. L.; Mensh, B.; Dudman, J. T.; Nelson, S. B.; Hantman, A. W., A repeated molecular architecture across thalamic pathways. *Nat. Neurosci.* **2019,** *22* (11), 1925-1935.

84. Taylor, C. F.; Paton, N. W.; Lilley, K. S.; Binz, P.-A.; Julian, R. K.; Jones, A. R.; Zhu, W.; Apweiler, R.; Aebersold, R.; Deutsch, E. W.; Dunn, M. J.; Heck, A. J. R.; Leitner, A.; Macht, M.; Mann, M.; Martens, L.; Neubert, T. A.; Patterson, S. D.; Ping, P.; Seymour, S. L.; Souda, P.; Tsugita, A.; Vandekerckhove, J.; Vondriska, T. M.; Whitelegge, J. P.; Wilkins, M. R.; Xenarios, I.; Yates, J. R.; Hermjakob, H., The minimum information about a proteomics experiment (MIAPE). *Nat. Biotechnol.* **2007,** *25* (8), 887-893.

85. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R., Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007,** *3* (3), 211-221.

86. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A., The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019,** *47* (D1), D442-D450.

87. Sud, M.; Fahy, E.; Cotter, D.; Azam, K.; Vadivelu, I.; Burant, C.; Edison, A.; Fiehn, O.; Higashi, R.; Nair, S. K.; Sumner, S.; Subramaniam, S., Metabolomics workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Res.* **2016,** *44* (D1), D463-D470.

88. Duncan, K. D.; Fyrestam, J.; Lanekoff, I., Advances in mass spectrometry based single-cell metabolomics. *Analyst* **2019,** *144* (3), 782-793.

89. Yin, L.; Zhang, Z.; Liu, Y.; Gao, Y.; Gu, J., Recent advances in single-cell analysis by mass spectrometry. *Analyst* **2019,** *144* (3), 824-845.

90. Comi, T. J.; Do, T. D.; Rubakhin, S. S.; Sweedler, J. V., Categorizing cells on the basis of their chemical profiles: progress in single-cell mass spectrometry. *J. Am. Chem. Soc.* **2017,** *139* (11), 3920-3929.

91. Couvillion, S. P.; Zhu, Y.; Nagy, G.; Adkins, J. N.; Ansong, C.; Renslow, R. S.; Piehowski, P. D.; Ibrahim, Y. M.; Kelly, R. T.; Metz, T. O., New mass spectrometry technologies contributing towards comprehensive and high throughput omics analyses of single cells. *Analyst* **2019,** *144* (3), 794-807.

92. Yang, Y.; Huang, Y.; Wu, J.; Liu, N.; Deng, J.; Luan, T., Single-cell analysis by ambient mass spectrometry. *Trends Anal. Chem.* **2017,** *90*, 14-26.

93. Qi, M.; Philip, M. C.; Yang, N.; Sweedler, J. V., Single cell neurometabolomics. *ACS Chem. Neurosci.* **2018,** *9* (1), 40-50.

94. Rubakhin, S. S.; Romanova, E. V.; Nemes, P.; Sweedler, J. V., Profiling metabolites and peptides in single cells. *Nat. Methods* **2011,** *8* (4), S20-S29.

95. Evers, T. M. J.; Hochane, M.; Tans, S. J.; Heeren, R. M. A.; Semrau, S.; Nemes, P.; Mashaghi, A., Deciphering metabolic heterogeneity by single-cell analysis. *Anal. Chem.* **2019,** *91* (21), 13314-13323.

96. Armbrecht, L.; Dittrich, P. S., Recent advances in the analysis of single cells. *Anal. Chem.* **2017,** *89* (1), 2-21.

97. Lombard-Banek, C.; Moody, S. A.; Nemes, P., High-sensitivity mass spectrometry for probing gene translation in single embryonic cells in the early frog (*Xenopus*) embryo. *Front. Cell Dev. Biol.* **2016,** *4*.

98. Lombard-Banek, C.; Portero, E. P.; Onjiko, R. M.; Nemes, P., New-generation mass spectrometry expands the toolbox of cell and developmental biology. *Genesis* **2017,** *55* (1-2), 1-14.

99. Onjiko, R. M.; Portero, E. P.; Nemes, P., Single-cell metabolomics with capillary electrophoresis–mass spectrometry. In *Capillary Electrophoresis–Mass Spectrometry for Metabolomics*, Ramautar, R., Ed. The Royal Society of Chemistry: **2018**; pp 209-224.

100. Sun, L.; Dubiak, K. M.; Peuchen, E. H.; Zhang, Z.; Zhu, G.; Huber, P. W.; Dovichi, N. J., Single cell proteomics using frog (*Xenopus laevis*) blastomeres isolated from early stage embryos, which form a geometric progression in protein content. *Anal. Chem.* **2016,** *88* (13), 6653-6657.

101. Saha-Shah, A.; Esmaeili, M.; Sidoli, S.; Hwang, H.; Yang, J.; Klein, P. S.; Garcia, B. A., Single cell proteomics by data-independent acquisition to study embryonic asymmetry in *Xenopus laevis*. *Anal. Chem.* **2019,** *91* (14), 8891-8899.

102. Budnik, B.; Levy, E.; Harmange, G.; Slavov, N., SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **2018,** *19* (1), 1-12.

103. Specht, H.; Emmott, E.; Petelski, A. A.; Huffman, G.; Perlman, D.; Serra, M.; Kharchenko, P.; Koller, A.; Slavov, N., Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity. *bioRxiv* **2020,** *665307*.

104. Cahill, J. F.; Riba, J.; Kertesz, V., Rapid, untargeted chemical profiling of single cells in their native environment. *Anal. Chem.* **2019,** *91* (9), 6118-6126.

105. Cong, Y. Z.; Liang, Y. R.; Motamedchaboki, K.; Huguet, R.; Truong, T.; Zhao, R.; Shen, Y. F.; Lopez-Ferrer, D.; Zhu, Y.; Kelly, R. T., Improved single-cell proteome coverage using narrow-bore packed nanoLC columns and ultrasensitive mass spectrometry. *Anal. Chem.* **2020,** *92* (3), 2665-2671.

106. Wuhr, M.; Guttler, T.; Peshkin, L.; McAlister, G. C.; Sonnett, M.; Ishihara, K.; Groen, A. C.; Presler, M.; Erickson, B. K.; Mitchison, T. J.; Kirschner, M. W.; Gygi, S. P., The nuclear proteome of a vertebrate. *Curr. Biol.* **2015,** *25* (20), 2663-2671.

107. Ivanov, A. R.; Zang, L.; Karger, B. L., Low-attomole electrospray ionization MS and MS/MS analysis of protein tryptic digests using 20-μm-i.d. polystyrene−divinylbenzene monolithic capillary columns. *Anal. Chem.* **2003,** *75* (20), 5306-5316.

108. Zhu, Y.; Piehowski, P. D.; Zhao, R.; Chen, J.; Shen, Y. F.; Moore, R. J.; Shukla, A. K.; Petyuk, V. A.; Campbell-Thompson, M.; Mathews, C. E.; Smith, R. D.; Qian, W. J.; Kelly, R. T., Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* **2018,** *9* (1), 1-10.

109. Osbourn, D. M.; Weiss, D. J.; Lunte, C. E., On-line preconcentration methods for capillary electrophoresis. *Electrophoresis* **2000,** *21* (14), 2768-2779.

110. Drouin, N.; van Mever, M.; Zhang, W.; Tobolkina, E.; Ferre, S.; Servais, A.-C.; Gou, M.-J.; Nyssen, L.; Fillet, M.; Lageveen-Kammeijer, G. S. M.; Nouta, J.; Chetwynd, A. J.; Lynch, I.; Thorn, J. A.; Meixner, J.; Lößner, C.; Taverna, M.; Liu, S.; Tran, N. T.; Francois, Y.; Lechner, A.; Nehmé, R.; Al Hamoui Dit Banni, G.; Nasreddine, R.; Colas, C.; Lindner, H. H.; Faserl, K.; Neusüß, C.; Nelke, M.; Lämmerer, S.; Perrin, C.; Bich-Muracciole, C.; Barbas, C.; Gonzálvez, Á. L.; Guttman, A.; Szigeti, M.; Britz-McKibbin, P.; Kroezen, Z.; Shanmuganathan, M.; Nemes, P.; Portero, E. P.; Hankemeier, T.; Codesido, S.; González-Ruiz, V.; Rudaz, S.; Ramautar, R., Capillary electrophoresis-mass spectrometry at trial by metabo-ring: effective electrophoretic mobility for reproducible and robust compound annotation. *Anal. Chem.* **2020,** *92* (20), 14103-14112.

111. Ramautar, R.; Somsen, G. W.; de Jong, G. J., CE-MS for metabolomics: developments and applications in the period 2016–2018. *Electrophoresis* **2019,** *40* (1), 165-179.

112. Ramautar, R.; Somsen, G. W.; de Jong, G. J., CE–MS for metabolomics: developments and applications in the period 2014–2016. *Electrophoresis* **2017,** *38* (1), 190-202.

113. Liao, H.-W.; Rubakhin, S. S.; Philip, M. C.; Sweedler, J. V., Enhanced single-cell metabolomics by capillary electrophoresis electrospray ionization-mass spectrometry with field amplified sample injection. *Anal. Chim. Acta* **2020,** *1118*, 36-43.

114. Choi, S. B.; Lombard-Banek, C.; Muñoz-Llancao, P.; Manzini, M. C.; Nemes, P., Enhanced peptide detection toward single-neuron proteomics by reversed-phase fractionation capillary electrophoresis mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2018,** *29* (5), 913-922.

115. Kawai, T.; Ota, N.; Okada, K.; Imasato, A.; Owa, Y.; Morita, M.; Tada, M.; Tanaka, Y., Ultrasensitive single cell metabolomics by capillary electrophoresis–mass spectrometry with a thin-walled tapered emitter and large-volume dual sample preconcentration. *Anal. Chem.* **2019,** *91* (16), 10564-10572.

116. Portero, E. P.; Nemes, P., Dual cationic–anionic profiling of metabolites in a single identified cell in a live *Xenopus laevis* embryo by microprobe CE-ESI-MS. *Analyst* **2019,** *144* (3), 892-900.
117. Lombard-Banek, C.; Yu, Z.; Swiercz, A. P.; Marvar, P. J.; Nemes, P., A microanalytical capillary electrophoresis mass spectrometry assay for quantifying angiotensin peptides in the brain. *Anal. Bioanal. Chem.* **2019,** *411* (19), 4661-4671.
118. Choi, S. B.; Munoz-Llancao, P.; Manzini, M. C.; Nemes, P., Data-dependent acquisition ladder for capillary electrophoresis mass spectrometry-based ultrasensitive (neuro)proteomics. *Anal. Chem.* **2021,** *93* (48), 15964-15972.
119. Choi, S. B.; Polter, A. M.; Nemes, P. In *Mass spectrometry-based proteomics of single dopaminergic mouse neurons identified by whole-cell electrophysiology*, Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, March 17; Philadelphia, PA **2019**.
120. Choi, S. B.; Polter, A. M.; Nemes, P. In *Single-cell mass spectrometry for proteomic analysis of patch-clamp electrophysiology identified dopaminergic neurons*, United States Human Proteome Organization, March 3; Washington DC, **2019**.
121. Onjiko, R. M.; Plotnick, D. O.; Moody, S. A.; Nemes, P., Metabolic comparison of dorsal *versus* ventral cells directly in the live 8-cell frog embryo by microprobe single-cell CE-ESI-MS. *Anal. Methods* **2017,** *9* (34), 4964-4970.
122. Onjiko, R. M.; Morris, S. E.; Moody, S. A.; Nemes, P., Single-cell mass spectrometry with multi-solvent extraction identifies metabolic differences between left and right blastomeres in the 8-cell frog (*Xenopus*) embryo. *Analyst* **2016,** *141* (12), 3648-3656.
123. Onjiko, R. M.; Nemes, P.; Moody, S. A., Altering metabolite distribution at *Xenopus* cleavage stages affects left-right gene expression asymmetries. *Genesis* **2021,** *59* (5-6), 1-8.
124. Portero, E. P.; Andrews, A. J.; Nemes, P. In *Probing metabolic pathways during early embryonic development using stable isotope labeling and single-cell mass spectrometry*, ASMS Conference on Mass Spectrometry and Allied Topics, June 6; Atlanta, GA, **2019**.
125. Portero, E. P.; Pade, L.; Nemes, P. In *Single-cell mass spectrometry reveals cell-to-cell communication in the same vertebrate (Xenopus laevis) frog embryo*, ASMS Conference on Mass Spectrometry and Allied Topics, June 12; Online, **2020**.

126. Aerts, J. T.; Louis, K. R.; Crandall, S. R.; Govindaiah, G.; Cox, C. L.; Sweedler, J. V., Patch clamp electrophysiology and capillary electrophoresis-mass spectrometry metabolomics for single cell characterization. *Anal. Chem.* **2014,** *86* (6), 3203-8.

127. Choi, S. B.; Polter, A. M.; Nemes, P. In *Integration of patch-clamp electrophysiology with single-cell mass spectrometry for proteomic analysis to extend the bioanalytical toolbox of neuroscience*, ASMS Conference on Mass Spectrometry and Allied Topics, June 4; Online, **2020**.

128. Moody, S. A., Microinjection of mRNAs and oligonucleotides. *Cold Spring Harbor protocols* **2018,** *2018* (12), 923-932.

129. Sive, H. L.; Grainger, R. M.; Harland, R. M., *Early development of Xenopus laevis: a laboratory manual*. Cold Spring Harbor Laboratory Press: **2000**.

130. Kaech, S.; Banker, G., Culturing hippocampal neurons. *Nat. Protoc.* **2006,** *1* (5), 2406-2415.

131. Klein, S. L., The 1st cleavage furrow demarcates the dorsal ventral axis in *Xenopus* embryos. *Dev. Biol.* **1987,** *120* (1), 299-304.

132. Polter, A. M.; Bishop, R. A.; Briand, L. A.; Graziane, N. M.; Pierce, R. C.; Kauer, J. A., Poststress block of kappa opioid receptors rescues long-term potentiation of inhibitory synapses and prevents reinstatement of cocaine seeking. *Biological Psychiatry* **2014,** *76* (10), 785-793.

133. Liu, Z. C.; Portero, E. P.; Jian, Y. R.; Zhao, Y. J.; Onjiko, R. M.; Zeng, C.; Nemes, P., Trace, machine learning of signal images for trace-sensitive mass spectrometry: a case study from single-cell metabolomics. *Anal. Chem.* **2019,** *91* (9), 5768-5776.

134. Xie, Y. R.; Castro, D. C.; Bell, S. E.; Rubakhin, S. S.; Sweedler, J. V., Single-cell classification using mass spectrometry through interpretable machine learning. *Anal. Chem.* **2020,** *92* (13), 9338-9347.

135. Sun, L.; Zhu, G.; Zhang, Z.; Mou, S.; Dovichi, N. J., Third-generation electrokinetically pumped sheath-flow nanospray interface with improved stability and sensitivity for automated capillary zone electrophoresis-mass spectrometry analysis of complex proteome digests. *J. Proteome Res.* **2015,** *14* (5), 2312-2321.

136. Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; Wolan,

D. W.; Spilker, M. E.; Benton, H. P.; Siuzdak, G., METLIN: a technology platform for identifying knowns and unknowns. *Anal. Chem.* **2018,** *90* (5), 3156-3164.

137. Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A., HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **2018,** *46* (D1), D608-D617.

138. Bateman, A.; Martin, M. J.; Orchard, S.; Magrane, M.; Agivetova, R.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E. H.; Britto, R.; Bursteinas, B.; Bye-A-Jee, H.; Coetzee, R.; Cukura, A.; Da Silva, A.; Denny, P.; Dogan, T.; Ebenezer, T.; Fan, J.; Castro, L. G.; Garmiri, P.; Georghiou, G.; Gonzales, L.; Hatton-Ellis, E.; Hussein, A.; Ignatchenko, A.; Insana, G.; Ishtiaq, R.; Jokinen, P.; Joshi, V.; Jyothi, D.; Lock, A.; Lopez, R.; Luciani, A.; Luo, J.; Lussi, Y.; Mac-Dougall, A.; Madeira, F.; Mahmoudy, M.; Menchi, M.; Mishra, A.; Moulang, K.; Nightingale, A.; Oliveira, C. S.; Pundir, S.; Qi, G. Y.; Raj, S.; Rice, D.; Lopez, M. R.; Saidi, R.; Sampson, J.; Sawford, T.; Speretta, E.; Turner, E.; Tyagi, N.; Vasudev, P.; Volynkin, V.; Warner, K.; Watkins, X.; Zaru, R.; Zellner, H.; Bridge, A.; Poux, S.; Redaschi, N.; Aimo, L.; Argoud-Puy, G.; Auchincloss, A.; Axelsen, K.; Bansal, P.; Baratin, D.; Blatter, M. C.; Bolleman, J.; Boutet, E.; Breuza, L.; Casals-Casas, C.; de Castro, E.; Echioukh, K. C.; Coudert, E.; Cuche, B.; Doche, M.; Dornevil, D.; Estreicher, A.; Famiglietti, M. L.; Feuermann, M.; Gasteiger, E.; Gehant, S.; Gerritsen, V.; Gos, A.; Gruaz-Gumowski, N.; Hinz, U.; Hulo, C.; Hyka-Nouspikel, N.; Jungo, F.; Keller, G.; Kerhornou, A.; Lara, V.; Le Mercier, P.; Lieberherr, D.; Lombardot, T.; Martin, X.; Masson, P.; Morgat, A.; Neto, T. B.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pourcel, L.; Pozzato, M.; Pruess, M.; Rivoire, C.; Sigrist, C.; Sonesson, K.; Stutz, A.; Sundaram, S.; Tognolli, M.; Verbregue, L.; Wu, C. H.; Arighi, C. N.; Arminski, L.; Chen, C. M.; Chen, Y. X.; Garavelli, J. S.; Huang, H. Z.; Laiho, K.; McGarvey, P.; Natale, D. A.; Ross, K.; Vinayaka, C. R.; Wang, Q. H.; Wang, Y. Q.; Yeh, L. S.; Zhang, J.; UniProt, C., UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021,** *49* (D1), D480-D489.

139. Karimi, K.; Fortriede, J. D.; Lotay, V. S.; Burns, K. A.; Wang, D. Z.; Fisher, M. E.; Pells, T. J.; James-Zorn, C.; Wang, Y.; Ponferrada, V. G.; Chu, S.; Chaturvedi, P.; Zorn, A. M.; Vize, P. D., Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res.* **2018,** *46* (D1), D861-D868.

140. Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **2013,** *113* (4), 2343-2394.

141. Moody, S. A., Cell lineage analysis in *Xenopus* embryos. *Methods in Molecular Biology* **2000,** *135*, 331-347.

142. Moody, S. A., Lineage tracing and fate mapping in *Xenopus* embryos. *Cold Spring Harbor protocols* **2018,** (12).

143. Viczian, A. S.; Zuber, M. E., A simple behavioral assay for testing visual function in *Xenopus laevis*. *J Vis Exp* **2014,** (88), e51726.

144. Ting, J. T.; Daigle, T. L.; Chen, Q.; Feng, G. P., Acute brain slice methods for adult and aging animals: application of targeted patch clamp analysis and optogenetics. In *Patch-Clamp Methods and Protocols, 2nd Edition*, Martina, M.; Taverna, S., Eds. **2014**; Vol. 1183, pp 221-242.

145. Grace, A. A.; Onn, S. P., Morphology and electrophysiological properties of immunocytochemically identified rat dopamine neurons recorded *in vitro*. *J. Neurosci.* **1989,** *9* (10), 3463-3481.

146. Khokha, M. K.; Chung, C.; Bustamante, E. L.; Gaw, L. W.; Trott, K. A.; Yeh, J.; Lim, N.; Lin, J. C.; Taverner, N.; Amaya, E.; Papalopulu, N.; Smith, J. C.; Zorn, A. M.; Harland, R. M.; Grammer, T. C., Techniques and probes for the study of *Xenopus tropicalis* development. *Dev. Dyn.* **2002,** *225* (4), 499-510.

147. Wallingford, J. B., Preparation of fixed *Xenopus* embryos for confocal imaging. *Cold Spring Harbor protocols* **2010,** (5), 1-7.

148. Kelly, R. T., Single-cell proteomics: progress and prospects. *Mol. Cell. Proteomics* **2020,** *19* (11), 1739-1748.

149. Levy, E.; Slavov, N., Single cell protein analysis for systems biology. In *Systems Biology*, Kolch, W.; Fey, D.; Ryan, C. J., Eds. Portland Press Ltd: London, **2018**; Vol. 62, pp 595-605.

150. Specht, H.; Slavov, N., Transformative opportunities for single-cell proteomics. *J. Proteome Res.* **2018,** *17* (8), 2565-2571.

151. Virant-Klun, I.; Leicht, S.; Hughes, C.; Krijgsveld, J., Identification of maturation-specific proteins by single-cell proteomics of human oocytes. *Mol. Cell. Proteomics* **2016,** *15* (8), 2616-2627.

152. Zhu, Y.; Clair, G.; Chrisler, W. B.; Shen, Y.; Zhao, R.; Shukla, A. K.; Moore, R. J.; Misra, R. S.; Pryhuber, G. S.; Smith, R. D.; Ansong, C.; Kelly, R. T., Proteomic analysis of single mammalian cells enabled by microfluidic nanodroplet sample preparation and ultrasensitive nanoLC‐MS. *Angew. Chem. Int. Edit.* **2018,** *57* (38), 12370-12374.

153. Li, S.; Plouffe, B. D.; Belov, A. M.; Ray, S.; Wang, X.; Murthy, S. K.; Karger, B. L.; Ivanov, A. R., An integrated platform for isolation, processing, and mass spectrometry-based proteomic profiling of rare cells in whole blood. *Mol. Cell. Proteomics* **2015,** *14* (6), 1672-1683.

154. Specht, H.; Emmott, E.; Petelski, A. A.; Huffman, R. G.; Perlman, D. H.; Serra, M.; Kharchenko, P.; Koller, A.; Slavov, N., Single-cell proteomic and transcriptomicanalysis of macrophage heterogeneity. *bioRxiv* **2020**.

155. Do, T. D.; Ellis, J. F.; Neumann, E. K.; Comi, T. J.; Tillmaand, E. G.; Lenhart, A. E.; Rubakhin, S. S.; Sweedler, J. V., Optically guided single cell mass spectrometry of rat dorsal root ganglia to profile lipids, peptides and proteins. *ChemPhysChem* **2018,** *19* (10), 1180-1191.

156. Lapainis, T.; Rubakhin, S. S.; Sweedler, J. V., Capillary electrophoresis with electrospray ionization mass spectrometric detection for single-cell metabolomics. *Anal. Chem.* **2009,** *81* (14), 5858-5864.

157. Macaulay, I. C.; Ponting, C. P.; Voet, T., Single-cell multiomics: multiple measurements from single cells. *Trends Genet.* **2017,** *33* (2), 155-168.

158. Moody, S. A.; Kline, M. J., Segregation of fate during cleavage of frog (*Xenopus laevis*) blastomeres. *Anat. Embryol.* **1990,** *182* (4), 347-362.

159. Sun, L.; Zhu, G.; Zhao, Y.; Yan, X.; Mou, S.; Dovichi, N. J., Ultrasensitive and fast bottom-up analysis of femtogram amounts of complex proteome digests. *Angew. Chem. Int. Edit.* **2013,** *52* (51), 13661-13664.

160. Xia, J. G.; Wishart, D. S., Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* **2011,** *6* (6), 743-760.

161. Xue, J.; Guijas, C.; Benton, H. P.; Warth, B.; Siuzdak, G., METLIN MS^2 molecular standards database: a broad chemical and biological resource. *Nat. Methods* **2020,** *17* (11), 953–954.

162. Wühr, M.; Freeman, R. M.; Presler, M.; Horb, M. E.; Peshkin, L.; Gygi, S. P.; Kirschner, M. W., Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. *Curr. Biol.* **2014,** *24* (13), 1467-1475.

163. Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **2014,** *13* (9), 2513-2526.

164. Hirose, G.; Jacobson, M., Clonal organization of the central nervous system of the frog: I. Clones stemming from individual blastomeres of the 16-cell and earlier stages. *Dev. Biol.* **1979,** *71* (2), 191-202.

165. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K., KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017,** *45* (D1), D353-D361.

166. De Domenico, E.; Owens, N. D. L.; Grant, I. M.; Gomes-Faria, R.; Gilchrist, M. J., Molecular asymmetry in the 8-cell stage *Xenopus tropicalis* embryo described by single blastomere transcript sequencing. *Dev. Biol.* **2015,** *408* (2), 252-268.

167. Henry, J. J.; Tsonis, P. A., Molecular and cellular aspects of amphibian lens regeneration. *Prog. Retin. Eye Res.* **2010,** *29* (6), 543-555.

168. Tatapudy, S.; Aloisio, F.; Barber, D.; Nystul, T., Cell fate decisions: emerging roles for metabolic signals and cell morphology. *EMBO Rep.* **2017,** *18* (12), 2105-2118.

169. Mizuno, H.; Tsuyama, N.; Harada, T.; Masujima, T., Live single-cell video-mass spectrometry for cellular and subcellular molecular detection and cell classification. *J. Mass Spectrom.* **2008,** *43* (12), 1692-1700.

170. Plouffe, B. D.; Murthy, S. K.; Lewis, L. H., Fundamentals and application of magnetic particles in cell isolation and enrichment: a review. *Rep. Prog. Phys.* **2015,** *78* (016601), 1-38.

171. Bonner, W. A.; Sweet, R. G.; Hulett, H. R.; Herzenberg, L. A., Fluorescence activated cell sorting. *Rev. Sci. Instrum.* **1972,** *43* (3), 404-409.

172. Liu, L.; Cheung, T. H.; Charville, G. W.; Rando, T. A., Isolation of skeletal muscle stem cells by fluorescence-activated cell sorting. *Nat. Protoc.* **2015,** *10* (10), 1612-1624.

173. Levenberg, S.; Ferreira, L. S.; Chen-Konak, L.; Kraehenbuehl, T. P.; Langer, R., Isolation, differentiation and characterization of vascular cells derived from human embryonic stem cells. *Nat. Protoc.* **2010,** *5* (6), 1115-1126.

174. Swennenhuis, J. F.; Reumers, J.; Thys, K.; Aerssens, J.; Terstappen, L., Efficiency of whole genome amplification of single circulating tumor cells enriched by CellSearch and sorted by FACS. *Genome Med.* **2013,** *5* (11), 1-11.

175. Ramazi, S.; Zahiri, J., Post-translational modifications in proteins: resources, tools and prediction methods. *Database* **2021**, 20.

176. Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M., Global quantification of mammalian gene expression control. *Nature* **2011,** *473* (7347), 337-342.

177. Zubarev, R. A., The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **2013,** *13* (5), 723-726.

178. Qian, W. J.; Kaleta, D. T.; Petritis, B. O.; Jiang, H. L.; Liu, T.; Zhang, X.; Mottaz, H. M.; Varnum, S. M.; Camp, D. G.; Huang, L.; Fang, X. M.; Zhang, W. W.; Smith, R. D., Enhanced detection of low abundance human plasma proteins using a tandem IgY12-SuperMix immunoaffinity separation strategy. *Mol. Cell. Proteomics* **2008,** *7* (10), 1963-1973.

179. Domon, B.; Aebersold, R., Mass spectrometry and protein analysis. *Science* **2006,** *312* (5771), 212-217.

180. Hu, L. H.; Ye, M. L.; Jiang, X. G.; Feng, S.; Zou, H. F., Advances in hyphenated analytical techniques for shotgun proteome and peptidome analysis - A review. *Anal. Chim. Acta* **2007,** *598* (2), 193-204.

181. Stadlmann, J.; Hudecz, O.; Krssakova, G.; Ctortecka, C.; Van Raemdonck, G.; Op De Beeck, J.; Desmet, G.; Penninger, J. M.; Jacobs, P.; Mechtler, K., Improved sensitivity in low-input proteomics using micropillar array-based chromatography. *Anal. Chem.* **2019,** *91* (22), 14203-14207.

182. Watson, D. G., A rough guide to metabolite identification using high resolution liquid chromatography mass spectrometry in metabolomic profiling in metazoans. *Computational and Structural Biotechnology Journal* **2013,** *4*, e201301005.

183. Blazenovic, I.; Kind, T.; Ji, J.; Fiehn, O., Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. *Metabolites* **2018,** *8* (31), 1-23.

184. Cao, Z. Y.; Yende, S.; Kellum, J. A.; Angus, D. C.; Robinson, R. A. S., Proteomics reveals age-related differences in the host immune response to sepsis. *J. Proteome Res.* **2014,** *13* (2), 422-432.

185. Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.; Bork, P.; Jensen, L. J.; Mering, C., STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019,** *47* (D1), D607-D613.

186. Mi, H. Y.; Lazareva-Ulitsky, B.; Loo, R.; Kejariwal, A.; Vandergriff, J.; Rabkin, S.; Guo, N.; Muruganujan, A.; Doremieux, O.; Campbell, M. J.; Kitano, H.; Thomas, P. D., The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res.* **2005,** *33*, D284-D288.

187. Strober, W., Trypan blue exclusion test of cell viability. *Current Protocols in Immunology* **2015,** *111*, A3.B.1-A3.B.3.

188. Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S. Z.; Bourque, G.; Wishart, D. S.; Xia, J. G., MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **2018,** *46* (W1), W486-W494.

189. Chakrabarty, R. P.; Chandel, N. S., Mitochondria as signaling organelles control mammalian stem cell fate. *Cell Stem Cell* **2021,** *28* (3), 394-408.

190. Gut, P.; Verdin, E., The nexus of chromatin regulation and intermediary metabolism. *Nature* **2013,** *502* (7472), 489-498.

191. Pettegrew, J. W.; Levine, J.; McClure, R. J., Acetyl-L-carnitine physical-chemical, metabolic, and therapeutic properties: relevance for its mode of action in Alzheimer's disease and geriatric depression. *Mol. Psychiatry* **2000,** *5* (6), 616-632.

192. Martinez, Y.; Li, X.; Liu, G.; Bin, P.; Yan, W. X.; Mas, D.; Valdivie, M.; Hu, C. A. A.; Ren, W. K.; Yin, Y. L., The role of methionine on metabolism, oxidative stress, and diseases. *Amino Acids* **2017,** *49* (12), 2091-2098.

193. Vaz, F. M.; Wanders, R. J. A., Carnitine biosynthesis in mammals. *Biochem. J.* **2002,** *361*, 417-429.

194. Ferreira, G. C.; McKenna, M. C., L-carnitine and acetyl-L-carnitine roles and neuroprotection in developing brain. *Neurochem. Res.* **2017,** *42* (6), 1661-1675.

195. Tucek, S., Regulation of acetylcholine synthesis in the brain. *J. Neurochem.* **1985,** *44* (1), 11-24.

196. Bottiglieri, T., S-adenosyl-L-methionine (SAMe): from the bench to the bedside-molecular basis of a pleiotrophic molecule. *Am. J. Clin. Nutr.* **2002,** *76* (5), 1151S-1157S.

197. Martinez-Reyes, I.; Chandel, N. S., Mitochondrial TCA cycle metabolites control physiology and disease. *Nat. Commun.* **2020,** *11* (1), 1-11.

198. Madiraju, P.; Pande, S. V.; Prentki, M.; Madiraju, S. R. M., Mitochondrial acetylcarnitine provides acetyl groups for nuclear histone acetylation. *Epigenetics* **2009,** *4* (6), 296-299.

199. Mews, P.; Donahue, G.; Drake, A. M.; Luczak, V.; Abel, T.; Berger, S. L., Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. *Nature* **2017,** *546* (7658), 381-386.

200. Sidoli, S.; Trefely, S.; Garcia, B. A.; Carrer, A., Integrated analysis of acetyl-CoA and histone modification via mass spectrometry to investigate metabolically driven acetylation. In *Cancer Metabolism: Methods and Protocols*, Haznadar, M., Ed. Humana Press Inc: Totowa, **2019**; Vol. 1928, pp 125-147.

201. Moussaieff, A.; Rouleau, M.; Kitsberg, D.; Cohen, M.; Levy, G.; Barasch, D.; Nemirovski, A.; Shen-Orr, S.; Laevsky, I.; Amit, M.; Bomze, D.; Elena-Herrmann, B.; Scherf, T.; Nissim-Rafinia, M.; Kempa, S.; Itskovitz-Eldor, J.; Meshorer, E.; Aberdam, D.; Nahmias, Y., Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab.* **2015,** *21* (3), 392-402.

202. Bond, M. R.; Hanover, J. A., O-GlcNAc cycling: a link between metabolism and chronic disease. In *Annual Review of Nutrition, Vol 33*, Cousins, R. J., Ed. Annual Reviews: Palo Alto, **2013**; Vol. 33, pp 205-229.

203. Sakabe, K.; Wang, Z. H.; Hart, G. W., β-*N*-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (46), 19915-19920.

204. Davie, J. R., Inhibition of histone deacetylase activity by butyrate. *J. Nutr.* **2003,** *133* (7), 2485S-2493S.

205. Chriett, S.; Dabek, A.; Wojtala, M.; Vidal, H.; Balcerczyk, A.; Pirola, L., Prominent action of butyrate over beta-hydroxybutyrate as histone deacetylase inhibitor, transcriptional modulator and anti-inflammatory molecule. *Sci Rep* **2019,** *9* (1), 1-14.

206. Yan, S.; Yang, X. F.; Liu, H. L.; Fu, N.; Ouyang, Y.; Qing, K., Long-chain acyl-CoA synthetase in fatty acid metabolism involved in liver and other diseases: an update. *World J. Gastroenterol.* **2015,** *21* (12), 3492-3498.

207. Abdelkreem, E.; Harijan, R. K.; Yamaguchi, S.; Wierenga, R. K.; Fukao, T., Mutation update on ACAT1 variants associated with mitochondrial acetoacetyl-CoA thiolase (T2) deficiency. *Hum. Mutat.* **2019,** *40* (10), 1641-1663.

208. Safran, M.; Dalah, I.; Alexander, J.; Rosen, N.; Stein, T. I.; Shmoish, M.; Nativ, N.; Bahir, I.; Doniger, T.; Krug, H.; Sirota-Madi, A.; Olender, T.; Golan, Y.; Stelzer, G.; Harel, A.; Lancet, D., GeneCards Version 3: the human gene integrator. *Database* **2010**, 16.

209. Miyazawa, H.; Aulehla, A., Revisiting the role of metabolism during development. *Development* **2018,** *145* (19), 1-11.

210. Otsuki, L.; Cheetham, S. W.; Brand, A. H., Freedom of expression: cell-type-specific gene profiling. *Wiley Interdiscip. Rev. Dev. Biol.* **2014,** *3* (6), 429-443.

211. Tang, F. C.; Lao, K. Q.; Surani, M. A., Development and applications of single-cell transcriptome analysis. *Nat. Methods* **2011,** *8* (4), S6-S11.

212. Bach, B. A.; Knape, W. A.; Edinger, M. G.; Tubbs, R. R., Improved sensitivity and resolution in the flow cytometric DNA analysis of human solid tumor specimens. Use of in vitro fine-needle aspiration and uniform staining reagents. *Am. J. Clin. Pathol.* **1991,** *96* (5), 615-627.

213. Morris, J.; Singh, J. M.; Eberwine, J. H., Transcriptome analysis of single cells. *J. Vis. Exp.* **2011,** (50), e2634.

214. EmmertBuck, M. R.; Bonner, R. F.; Smith, P. D.; Chuaqui, R. F.; Zhuang, Z. P.; Goldstein, S. R.; Weiss, R. A.; Liotta, L. A., Laser capture microdissection. *Science* **1996,** *274* (5289), 998-1001.

215. Espina, V.; Wulfkuhle, J. D.; Calvert, V. S.; VanMeter, A.; Zhou, W. D.; Coukos, G.; Geho, D. H.; Petricoin, E. F.; Liotta, L. A., Laser-capture microdissection. *Nat. Protoc.* **2006,** *1* (2), 586-603.

216. Gijs, M. A. M.; Lacharme, F.; Lehmann, U., Microfluidic applications of magnetic particles for biological analysis and catalysis. *Chem. Rev.* **2010,** *110* (3), 1518-1563.

217. Saliba, A. E.; Westermann, A. J.; Gorski, S. A.; Vogel, J., Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* **2014,** *42* (14), 8845-8860.

218. Fu, A. Y.; Spence, C.; Scherer, A.; Arnold, F. H.; Quake, S. R., A microfabricated fluorescence-activated cell sorter. *Nat. Biotechnol.* **1999,** *17* (11), 1109-1111.

219. Llufrio, E. M.; Wang, L. J.; Naser, F. J.; Patti, G. J., Sorting cells alters their redox state and cellular metabolome. *Redox Biol.* **2018,** *16*, 381-387.

220. Binek, A.; Rojo, D.; Godzien, J.; Ruperez, F. J.; Nunez, V.; Jorge, I.; Ricote, M.; Vazquez, J.; Barbas, C., Flow cytometry has a significant impact on the cellular metabolome. *J. Proteome Res.* **2019,** *18* (1), 169-181.

221. Miltenyi, S.; Muller, W.; Weichel, W.; Radbruch, A., High gradient magnetic cell separation with MACS. *Cytometry* **1990,** *11* (2), 231-238.

222. Han, K. H.; Frazier, A. B., Paramagnetic capture mode magnetophoretic microseparator for high efficiency blood cell separations. *Lab Chip* **2006,** *6* (2), 265-273.

223. Yang, Y.; Wang, H.; He, J.; Shi, W. C.; Jiang, Z. M.; Gao, L. N.; Jiang, Y.; Ni, R.; Yang, Q. F.; Luo, L. F., A single-cell-resolution fate map of endoderm reveals demarcation of pancreatic progenitors by cell cycle. *Proc. Natl. Acad. Sci. U. S. A.* **2021,** *118* (25), e2025793118.

224. Pang, Z. Q.; Chong, J.; Zhou, G. Y.; Morais, D. A. D.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P. E.; Li, S. Z.; Xia, J. G., MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **2021,** *49* (W1), W388-W396.

225. Berndt, F.; Shah, G.; Power, R. M.; Brugues, J.; Huisken, J., Dynamic and non-contact 3D sample rotation for microscopy. *Nat. Commun.* **2018,** *9* (1), 1-7.

226. Lewin, M.; Carlesso, N.; Tung, C. H.; Tang, X. W.; Cory, D.; Scadden, D. T.; Weissleder, R., Tat peptide-derivatized magnetic nanoparticles allow in vivo tracking and recovery of progenitor cells. *Nat. Biotechnol.* **2000,** *18* (4), 410-414.

227. Adams, J. D.; Kim, U.; Soh, H. T., Multitarget magnetic activated cell sorter. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (47), 18165-18170.

228. Aki, A.; Ito, O.; Morimoto, H.; Nagaoka, Y.; Nakajima, Y.; Mizuki, T.; Hanajiri, T.; Usami, R.; Maekawa, T., Capture of nonmagnetic particles and living cells using a microelectromagnetic system. *J. Appl. Phys.* **2008,** *104* (094509), 1-5.

229. Tenje, M.; Fornell, A.; Ohlin, M.; Nilsson, J., Particle manipulation methods in droplet microfluidics. *Anal. Chem.* **2018,** *90* (3), 1434-1443.

230. Zitka, O.; Skalickova, S.; Gumulec, J.; Masarik, M.; Adam, V.; Hubalek, J.; Trnkova, L.; Kruseova, J.; Eckschlager, T.; Kizek, R., Redox status expressed as GSH:GSSG ratio as a marker for oxidative stress in paediatric tumour patients. *Oncol. Lett.* **2012,** *4* (6), 1247-1253.

231. Hasegawa, Y.; Otoki, Y.; McClorry, S.; Coates, L. C.; Lombardi, R. L.; Taha, A. Y.; Slupsky, C. M., Optimization of a method for the simultaneous extraction of polar and non-polar oxylipin metabolites, DNA, RNA, small RNA, and protein from a single small tissue sample. *Methods Protoc.* **2020,** *3* (61), 1-14.

232. Chomczynski, P.; Sacchi, N., The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat. Protoc.* **2006,** *1* (2), 581-585.

233. Salem, M.; Bernach, M.; Bajdzienko, K.; Giavalisco, P., A simple fractionated extraction method for the comprehensive analysis of metabolites, lipids, and proteins from a single sample. *J. Vis. Exp.* **2017,** (124), e55802.

234. Vorreiter, F.; Richter, S.; Peter, M.; Baumann, S.; von Bergen, M.; Tomm, J. M., Comparison and optimization of methods for the simultaneous extraction of DNA, RNA, proteins, and metabolites. *Anal. Biochem.* **2016,** *508*, 25-33.