ABSTRACT

Title of Dissertation:	SPATIOTEMPORAL PROTEOMIC APPROACHES FOR INVESTIGATING PATTERNING DURING EMBRYONIC DEVELOPMENT
	Leena Rajendra Pade, Doctor of Philosophy, 2024
Dissertation directed by:	Prof. Peter Nemes, Department of Chemistry and Biochemistry

Characterization of molecular events as embryonic cells give rise to tissues and organs raises a potential to better understand normal development and design remedies for diseases. In this work, I integrated bioanalytical chemistry with neurodevelopmental biology to uncover mechanisms underlying tissue induction in a developing embryo. Specifically, I developed ultrasensitive proteomic approaches to study the remodeling of the proteome as embryonic cells differentiate in space and time to induce tissue formation. This dissertation discusses the design and development of proteomic strategies to deepen proteomic coverage from limited embryonic tissues. A novel sample preparation workflow and detection strategy was developed to address the challenge of interference from abundant proteins such as yolk in *Xenopus* tissues which in turn boosts the sensitivity of detecting low abundant proteins from complex limited amounts of tissues. The refined analytical workflow was implemented to study the development of critical signaling centers and stem cell populations and the tissues they induce to form in developing embryos.

Chapter 1 overviews the overarching goal of this research and introduces basic concepts of biology, research model, and key technologies that enable biological mass spectrometry analysis.

Chapter 2 describes development of an analytical approach to improve proteomic coverage from limited embryonic tissues.

Chapter 3 illustrates the development of sample preparation design tailored to process microscale sample volumes.

Chapter 4 probes into spatial proteomic heterogeneity that guides tissue differentiation during embryonic gastrulation using the above developed analytical methods.

Chapter 5 explores proteomic gradients present within spatially distinct branchial arches, which play a crucial role in craniofacial development.

Chapter 6 provides an overview of the significant analytical and biological achievements, along with a discussion on the potential implications and impacts of the newfound knowledge for future endeavors.

SPATIOTEMPORAL PROTEOMIC APPROACHES FOR INVESTIGATING PATTERNING DURING EMBRYONIC DEVELOPMENT

by

Leena Rajendra Pade

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
2024

Advisory Committee:
Professor Peter Nemes, Chair
Professor Neil Blough
Assistant Professor Myles B. Poulin
Professor Alice Mignerey
Professor Kan Cao, Dean's Representative

Dedication

In the loving memory of my father, whose relentless support and encouragement shaped the person I am today. Your passion for science and learning inspired me to excel in my academic endeavors. This dissertation is a tribute to your everlasting impact on my life.

Acknowledgements

I would like to express my sincere gratitude to all those who have contributed to the completion of this Ph.D. thesis. Their support, guidance, and encouragement have been invaluable throughout this challenging journey. First and foremost, I extend my deepest appreciation to my advisor, Dr. Peter Nemes, his confidence in my capability to steer my projects while providing constructive feedback to help me stay on track, has been instrumental in my academic development and shaping this research. Secondly, I would like to thank my advisor Dr. Ira Daar for providing me with the opportunity to be a part of his laboratory at the National Cancer Institute. I am deeply touched by his strong guidance and kind words that kept me motivated throughout my Ph.D. journey. I thank Dr. Jaeho Yoon, my mentor at NCI, for his guidance in facilitating my comprehension of the collaborative project with NCI.

I am grateful to the members of my dissertation committee, Dr. Neil Blough, Dr. Alice Mignerey, Dr. Myles Poulin, and Dr. Kan Cao for their time and insights on my thesis. Next, I would like to acknowledge the financial support provided by the National Cancer Institute, NIH, Cosmos scholarship, ASMS, and WCDG which enabled me to conduct the research, attend scientific conferences, and present my findings. I am grateful to the Graduate Office at the Department of Chemistry and Biochemistry for their indispensable support in helping me navigate the logistical and administrative details.

I would like to thank all the current members of Nemes Research Group including Bowen, Ailing, Fei, Ryan, Laura, Dash, Tarikul, and Yayra for sharing the responsibilities within the laboratory, help troubleshooting various research challenges, helping me prepare for my presentations, and striking fun conversations in the office. My time with the Nemes Research

Group has been more than just a professional journey; it's been a source of lasting friendships that have extended far beyond our lab. I'm thankful for the genuine connections formed with my colleagues, Dr. Kaitlyn Stepler and Dr. Kellen DeLaney. Your constant support and encouragement have meant the world to me. I am eternally grateful to Dr. Aparna Baxi, she has been a great mentor, colleague, friend, and constant source of inspiration. The moments we shared will forever hold a special place in my heart, and her absence is deeply felt. I extend my thanks to the Nemes Lab alumni, Dr. Jie Li, Dr. Sam Choi, Dr. Erika Portero, Dr. Camille Lombard, and Vi Quach, for their guidance and training during my early days in the lab.

My heartfelt appreciation goes to my family for their unwavering love, understanding, and encouragement throughout this academic endeavor. I would like to thank my mom, Anita Pade, and my brother, Sumeet Pade, for their support, which has been a constant source of strength and motivation during difficult times. My partner, Nilesh, has truly been my rock throughout my Ph.D. journey. He's been there for me, helping me tackle personal challenges and urging me to keep pushing forward. I owe so much to his unwavering support and encouragement; I couldn't have made it this far without him faithfully by my side. My beloved pet, Loki, has been a constant source of joy and helped me unwind after a strenuous day. I am indebted to my cousin Poorva who has been by my side through thick and thin, from the moment I set foot in the States. I also want to express my gratitude to my extended family members, who have supported me during various phases of my journey.

Lastly, I extend my gratitude to all my friends and colleagues who have provided emotional support and camaraderie, making this academic journey a more enriching and enjoyable experience. This thesis is the culmination of the collective efforts and support from various individuals and organizations, and I am sincerely thankful to each and every one of them.

Table of Contents

List of Tables

Appendix

List of Figures

Appendix

List of Abbreviations

MS	Mass Spectrometry
LC-MS	Liquid Chromatography mass spectrometry
HRMS	High resolution mass spectrometry
RPLC	Reversed-phase liquid chromatography
ESI	Electrospray ionization
DDA	Data dependent acquisition
CID	Collision induced dissociation
HCD	High energy collision induced dissociation
SPS	Synchronous precursor selection
TMT	Tandem mass tags
m/z	Mass to charge ratio
YODEC	Yolk depleted carrier
µYODE	Microscale yolk depletion
SMO	Spemann Mangold organizer
NE	Neural ectoderm
WE	Whole embryo
BA	Branchial arches
CNC	Cranial neural crest
BR	Biological replicates
TR	Technical replicates

Chapter 1: Introduction

(Parts of this chapter are reproduced and reused with permission from Pade, L. R., Stepler, K. E., Portero, E. P., DeLaney, K., & Nemes, P. (2024). Biological mass spectrometry enables spatiotemporal 'omics: From tissues to cells to organelles. Mass Spectrometry Reviews, 43(1), 106-138.)

Contribution: L.R.P. reviewed the literature, drafted the manuscript, and revised the contents.

Interdisciplinary efforts sculpted mass spectrometry (MS) into an indispensable tool in the molecular study of life processes. Virtually all levels of the molecular architecture—from genes to transcripts to proteins to peptides to metabolites—can now be measured using this technology, some in exquisite detail. MS detection and quantification of proteomes, peptidomes, and metabolomes ('omes) expanded the analytical toolbox of biology across the taxa, helping to generate and test new hypotheses. In this work, we harnessed the capabilities of MS to study critical processes during vertebrate embryonic development. The ability to detect thousands of proteins and metabolites from biological specimens provides a unique opportunity to explore the expression of molecules that govern cellular signaling and drive embryogenesis.

1.1 Biological question and research motivation

The overarching goal of this work is to uncover heterogeneity in proteomic expression within the embryo that regulates tissue induction and cell differentiation. For an embryo to develop normally, cells need to constantly communicate with each other to divide and form different types of tissues and organs. During gastrulation, such signaling is steered by a group of cells called the organizer within the embryo. In *Xenopus*, the Spemann-Mangold Organizer (SMO) is an important signaling center established during early development, it guides neighboring cells to form the neural tissue and it is responsible for dorsoanterior patterning during embryonic

development (**Fig. 1.1A**).[1] Identifying proteins produced in the SMO will provide insights into how embryonic cells are instructed to form neural tissue. To further understand tissue differentiation at later stages after the formation of three germ layers (endoderm, mesoderm, ectoderm), i.e., post gastrulation, we aim to study a subset of neural crest cells called cranial neural crest cells (CNCs). The neural crest cells are sometimes referred to as the fourth germ layer because of their importance. This group of multipotent progenitor cells is derived from the ectoderm and maintains their stemness at later stages. These cells migrate and differentiate into a variety of cell types and tissues such as osteocytes, pigment cells, connective tissue, sensory ganglia, glial cells, and smooth muscle tissue (**Fig. 1.1B**). CNCs specifically give rise to neural and non-neural tissues in the head and jaw region of vertebrates. Alterations during CNC development are known to cause severe congenital defects.[2] Identifying proteins during normal CNC development can help uncover mechanisms involved in the differentiation of stem cells into various tissue types.[3]

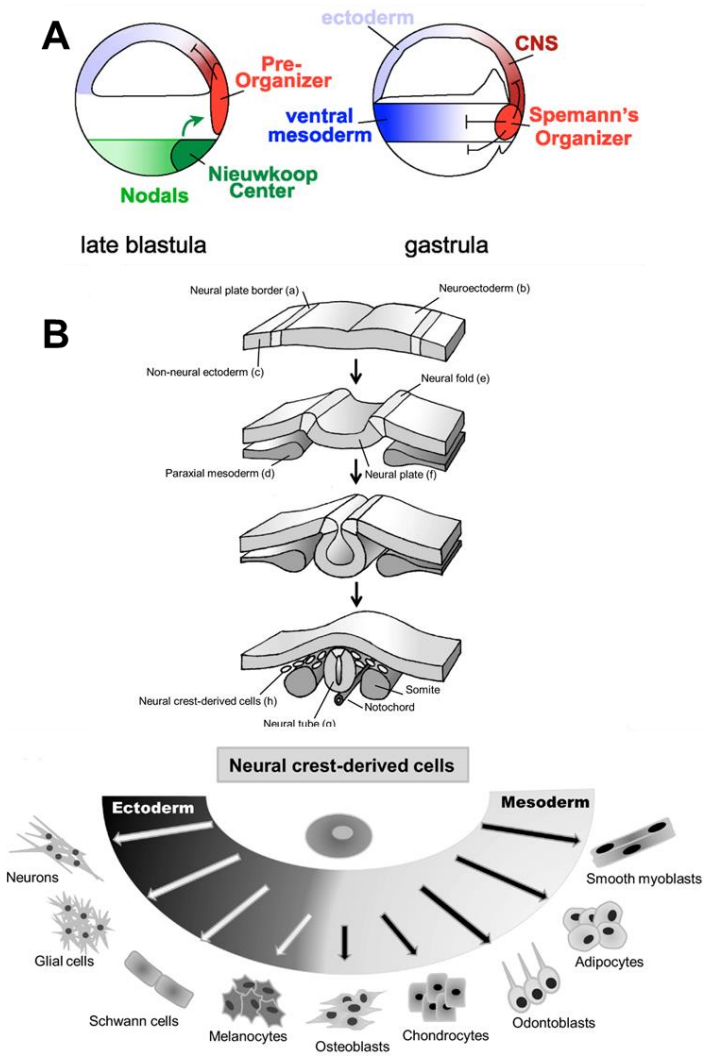

Figure 1.1. Key processes during embryonic development**. (A)** SMO establishing central nervous system (CNS) and ventral mesoderm in an early-stage embryo. **(B)** CNCs differentiate into different tissue types. Reproduced from references[4-6]

1.2 *Xenopus* as a biological model

Our group uses the South African clawed frog (*Xenopus laevis*), a vertebrate biological model, to study molecular processes during normal and impaired development.[7] **Table A1** highlights select publications that demonstrate the study of neurodevelopmental questions via proteomics using *Xenopus* across different scales (*see Appendix*). *Xenopus* offers several advantages that make it a valuable model to study embryonic development. It is an established tool to model human

diseases as it shares a close evolutionary history with and is genetically similar to humans. Over ~90% of human disease genes have homologs in *Xenopus* species. Females can be stimulated with gonadotropin hormone to lay eggs (500-1000) all year round. Embryos fertilize and develop externally, allowing close observation of early embryonic development. Furthermore, eggs are large (1.2 mm diameter) and have large embryonic cells (~250 µm). The large size of the embryonic cells with stereotypical pigmentation and cleavages permits easy identification and experimental manipulation such as injection of dyes and morpholinos. Another experimental advantage of the *Xenopus* model is the known cell fate of early embryonic cells. Reproducible fate maps are available for embryos at the 16- and 32-cell stage.[8-11] These maps can be used to trace different tissue lineages during development. Embryonic development is fast and temperature dependent[12] which allows the study of important biological events during embryogenesis such as mid-blastula transition, gastrulation, neurulation, and organogenesis in a short time frame. As my research aims involved studying tissue differentiation during early embryonic development, the amphibian species *Xenopus* proved to be an invaluable model as it enabled high throughput collection of tissues from important developmental stages and allowed timing proteomics experiments to developmental milestones.

1.3 MS Meets Research Needs

Life processes respond to intrinsic and extrinsic events with complex and dynamic molecular changes across broad spatial and temporal dimensions, and measuring these perturbations is key to their understanding. Systems biology promotes holistic understanding, in part, by characterizing all types of molecules and their potential interactions from organismal to subcellular scales. Vast signal amplification via polymerase chain reaction (PCR) facilitated next-generation sequencing of minute amounts of transcripts and genes, thus leading single-cell

sequencing to become quasi-routine. Subcellular RNA-seq, for example, recently uncovered transcriptomic differences between the cell body (soma) and dendrites in individual GABAergic and glutaminergic neurons during biological functions[13], raising the question of subcellular heterogeneity for functional molecules downstream of transcription. Without technologies capable of amplifying whole proteomes to metabolomes, information from sequencing is often used as a proxy for the proteome.

However, the relationship between mRNA expression, protein production, and metabolism is complex in dynamic systems[14]. Post-transcriptional and -translational regulation as well as protein turnover rates contribute to a widely variable correlation between transcription and translation. This relationship is further complicated by processes occurring heterogeneously in space and time, as was recently found during cell culturing[15], signal transduction between cells and neurons, differentiation of stem cells and developing organisms[16-17], cancer invasion and metastasis[18], and responses to drug treatments[19]. Similarly, bioactive peptides and metabolites, which also carry out important physiological functions, are the results of interconnected molecular pathways, complicating their predictions from information upstream. Elucidating these molecular relationships over space-time is important for understanding states of health and disease, but this goal requires direct detection of proteins, peptides, and metabolites.

MS, especially high-resolution MS (HRMS), emerged as the technology of choice to bridge this knowledge gap. It integrates exceptional molecular specificity with a capability for quantification, typically without the need for functional probes. Readers interested in the fundamentals of MS and HRMS, ranging from ion generation to mass analysis, are referred to recent reviews[20]. From tissues and cell cultures, hundreds to tens of thousands of biomolecules are routinely assessed in discovery studies with temporal resolution, with near-complete

coverage of proteomes[21-22]. For example, HRMS on two- (2-D) and three-dimensional (3-D) tissue cultures of human embryonic stem cells (hESCs) and induced pluripotent stem cells helped delineate molecular pathways critical for cell differentiation, cellular repair, and degeneration[23-25]. Quantitative HRMS on proliferating neuronal stem cells revealed proteomic alterations between cellular and secreted proteomes in response to the microenvironment that would be complicated to predict from the transcriptome[23]. HRMS quantified ~2,800 proteins and their temporal expression profiles at progressive stages of differentiation, offering a glimpse into molecular mechanisms toward developing next-generational cell-based therapies such as cell grafting[23]. Similarly, metabolic profiling of naïve and primed hESCs revealed alterations in epigenetic dynamics underpinning cell fate determination[24].

The advent of high-sensitive HRMS propelled ‘omics to dynamic and microscale biological systems. This group of technologies integrated existing approaches and invented new ones to accomplish trace-level sensitivity, sufficient to probe molecules at physiological concentrations. In doing so, they also allowed for scaling the spatial and temporal resolution of analysis to the specific biological systems and questions at focus. **Figure 1.2** conceptualizes broad spatiotemporal domains that contemporary HRMS is able to address. With sample collection capable of spatiotemporal scalability, high-sensitivity HRMS opened even more investigative possibilities. Rapid sampling enabled HRMS to assist studies on fast biochemical processes, including those underpinning cellular processes. For example, methods of sampling with duty cycles lasting minutes to seconds can collect cellular materials even within a ~15–45 min cell cycle in cleavage-stage embryos, as shown from the South African clawed frog (*Xenopus laevis*) in **Figure 1.2**. Sampling over minutes to hours brought molecular studies to important physiological processes, including the maintenance of circadian cycles, feeding, and response to

stress (see examples later). Additionally, scalability in physical dimensions raised the possibility of capturing spatial information on the distribution of molecules in specimens on the order of millimeters to microns. Mass spectrometry imaging (MSI) played an important role by enabling the mapping of proteins, peptides, and metabolites across biological tissues, supporting studies on exploring their functional roles, and aiding drug discovery. Profiling and imaging HRMS were recently adapted to subcellular structures. Chemical labeling enhanced spatial resolution to the domain of interacting molecules, thus vastly enhancing knowledge of the biochemistry underlying biology, such as signaling between neurons in the brain.

These various advancements expanded spatiotemporal HRMS to further support basic and translational studies. Next-generation separations, enhanced analyzers, and intelligent data acquisition strategies facilitated molecular identifications and quantification from minute amounts of materials. Thousands of proteins and metabolites can now be measured, even from single cells. For targeted proteins, heavy-element tagged antibodies extended HRMS to MSI and single-cell resolution. Novel mathematical algorithms helped process complex primary data in depth. User-friendly software packages and community-wide efforts to freely share them aided adoptability across laboratories needing sensitive ‘omics, raising an opportunity to support applications in basic biological and translational studies. Molecular information from these measurements holds answers to existing questions and supports new questions and hypotheses, which in turn require further advances in all aspects of technology and method development. Efforts in pursuit of sensitivity, scalability, and ‘omic coverage progressively expand our technological toolbox, thus accelerating the scientific cycle of exploration and discovery.

Over the last decades, technological advances broke down classical limitations in detection limits (sensitivity), speed, and scalability to enable the analysis of limited amounts of materials. This

chapter offers examples of innovations that extended MS-based ‘omics to study the spatiotemporal organization of diverse biological systems at the realms of the organism, organ, tissue, cell, and organelle. Innovations across all stages of the MS workflow (sample collection, preparation, detection, and data analysis), as highlighted here, have contributed to expanding the development of newer approaches tailored to address specific challenges faced in this work.

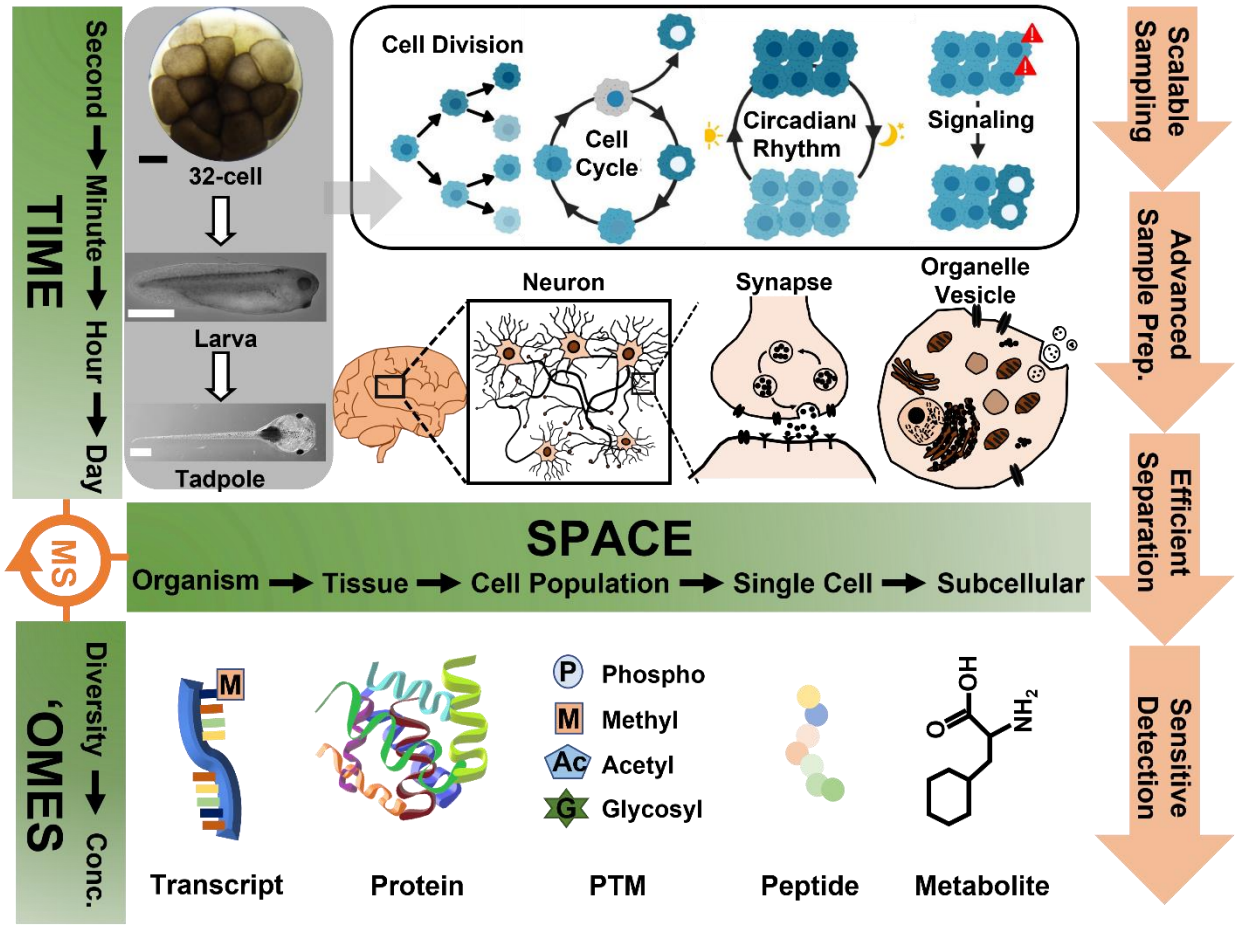

Figure 1.2. Understanding biological processes hinges on the analysis of molecular heterogeneity that they orchestrate over broad spatial and temporal domains. Recent technological innovations and the development of specialized methodologies in sample collection and preparation, chemical separation, and HRMS analysis and detection provide sufficient sensitivity and spatiotemporal resolution for studying molecular processes at the levels of transcriptomes, proteomes, peptidomes, and metabolomes (‘omes) in ever-deepening information content. Information from these studies fuels basic and translational research forward. Representative examples are highlighted from whole organismal to subcellular studies. Adapted with permission from[26-27].

1.4 General Workflow of MS-based ‘Omics

For spatiotemporal analysis, the interconnected steps of the analytical pipeline are presented in **Figure 1.3**. During sampling, specimens, such as tissues and cells, are collected and lysed to release their contents by a variety of methods and technologies, with each tailoring to specific benefits. The lysate is typically enriched for the biomolecules of interest, such as proteins or metabolites, and then subjected to various sample clean-up/purification steps to render compatibility for instrumental analysis. Various technologies are used for separation in solution, such as liquid chromatography (LC) and capillary electrophoresis (CE) or the gas phase, such as ion mobility spectrometry (IMS). The biomolecules are next ionized, detected, identified, and quantified by HRMS. An array of fragmentation (tandem/multistage MS) and analyzer-detector systems provide complementary speed and spectral resolution, ranging from quadrupoles to time-of-flight (TOF), orbitrap, and Fourier-transform ion cyclotron resonance (FTICR). The primary data from these measurements are processed and analyzed, usually with specialized MS bioinformatics software packages supporting identifications of proteins, peptides, metabolites, and their posttranslational modifications as well as relative or absolute quantification. Statistical and multivariate models from chemometrics allow for systematic evaluations of the observed effects between different experimental conditions under study. The results are interpreted in relation to canonical knowledge and the biological context. In successful investigations, the specifics of each step of the workflow are matched to the goals of the study, usually according to the following considerations (recall **Fig. 1.3**). The following sections focus on technologies that advanced spatiotemporal analysis in different biological specimens.

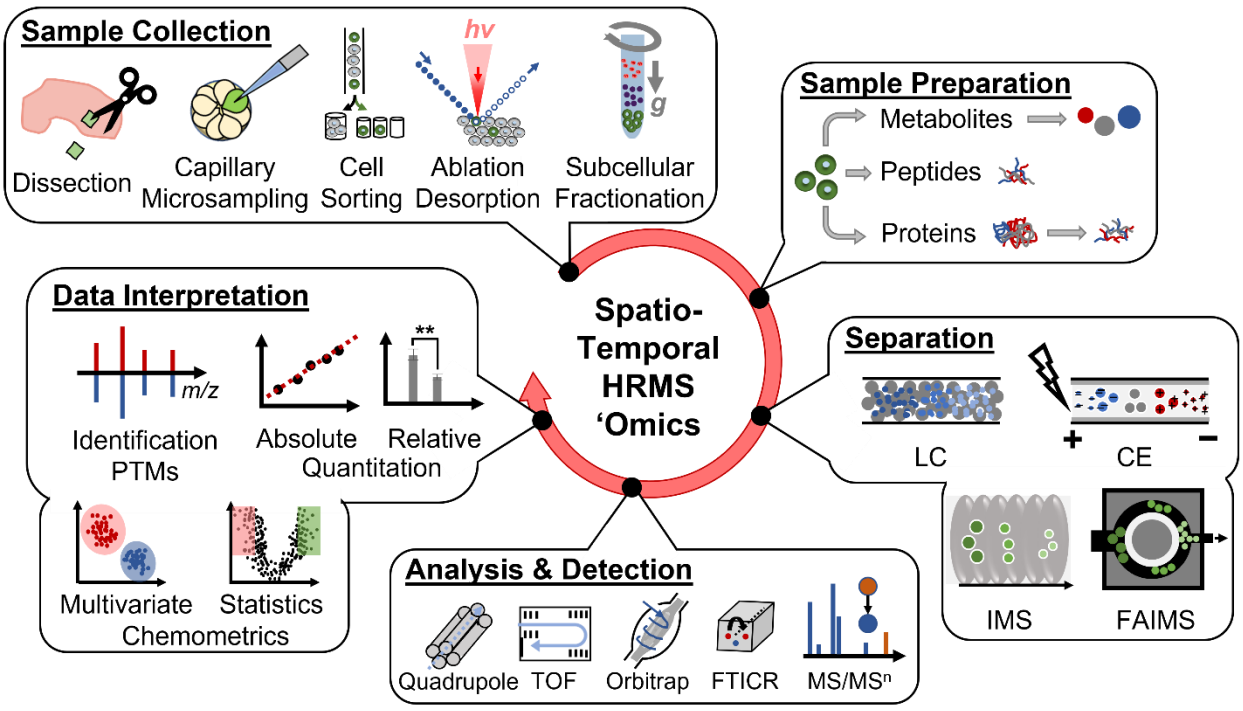

Figure 1.3. General experimental workflow and representative technologies enabling spatiotemporal MS. Scalable sample collection and processing allows for probing biological specimens across broad dimensions in space and time. Separation in the liquid and/or gas phase reduces spectral complexity. High-resolution mass analyzers, tandem/multistage fragmentation, and efficient detector systems support molecular identifications in high specificity and quantification in high sensitivity. Development of advanced tools from bioinformatics and chemometrics facilitates the analysis and interpretation of MS data with spatiotemporal insights. Technological abbreviations are in the text.

A. Sample Collection

The type of question and biological system at hand de facto determines the method of sampling. In tissues displaying heterogeneous biochemistry, such as brains and kidneys, local sampling with accuracy ensures contamination-free analysis from neighboring cell populations, thus aiding the interpretation of results. Likewise, sample collection with single-cell resolution unmasks differences between single cells (cell heterogeneity) within the same tissue, which are otherwise obscured due to signal averaging during cell pooling. As exemplified in **Figure 1.4**, various

strategies were developed to collect and purify tissue biopsies, small populations of cells, single cells, or subcellular fractions of interest with different speeds and operational conditions.

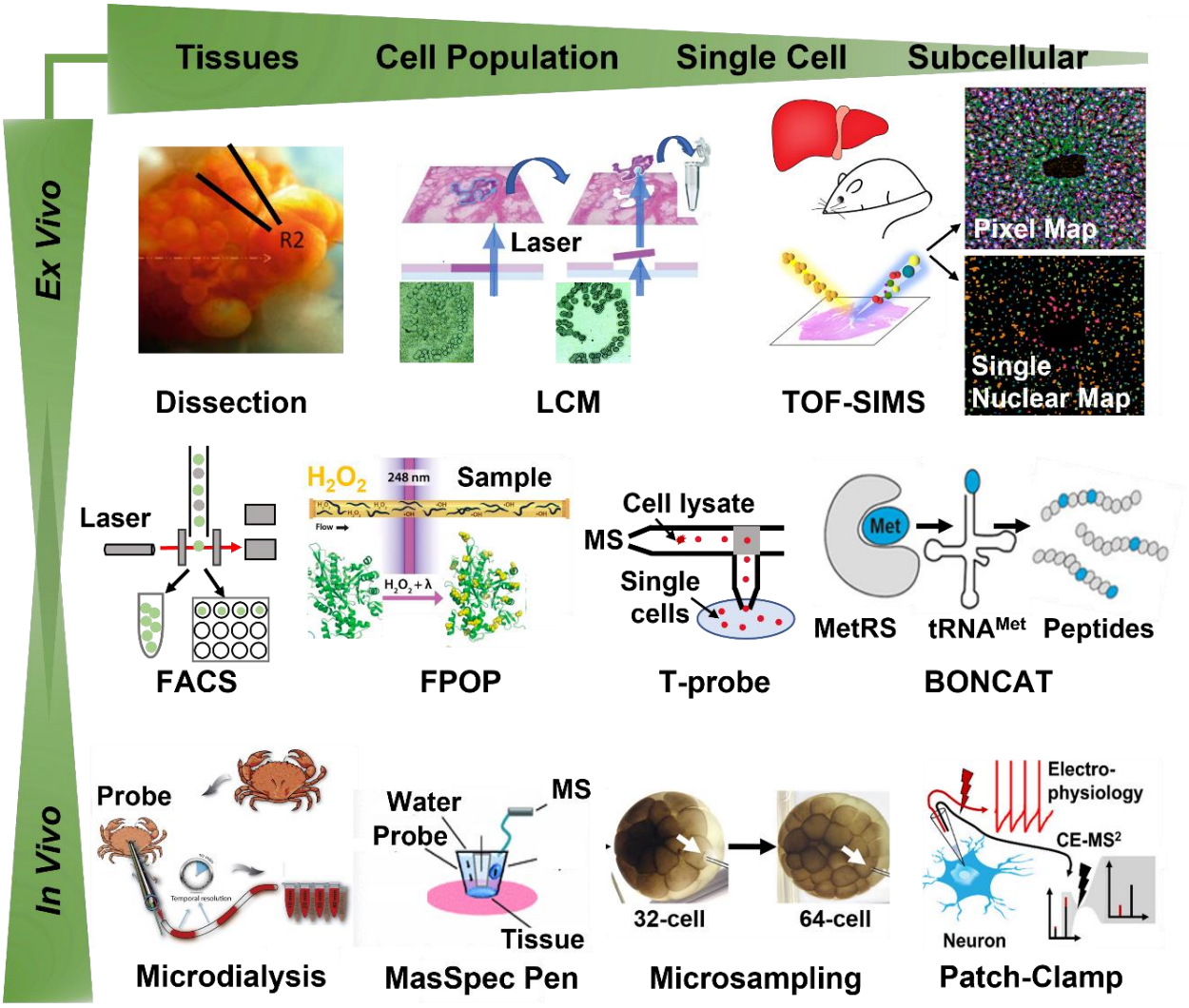

Figure 1.4. Representative strategies enabling scalable spatiotemporal sampling. Diverse technologies enabled HRMS with scalable operation to assess 'omics in broad types of biological systems and investigative contexts, ranging from ex vivo to in situ to in vivo analyses from small populations of cells to single cells to subcellular organelles. Technological abbreviations are in the text. Adapted with permission from[28-37].

i. From Biopsies to Single Cells

Manual dissection is simple and can have high accuracy. Surgical tools, ranging from sharp forceps to tungsten needles to eyebrow knives, are readily adaptable to excising distinct regions

of tissues. Dissection techniques, for example, enabled the isolation of spatially distinct colon tumor biopsies from humans[38], brain tissues from the mouse[39], large single cells[40-41] and small cell populations[42] from developing *X. laevis* embryos, neurons from *A. californica*[43], and distinct regions of nervous tissue from *Callinectes sapidus*[44]. Labeling with dyes or fluorescent reporters enhances the accuracy of dissections under a microscope. **Figure 1.4** demonstrates one such example, where identified cholinergic neurons from *A. californica* were dissected, permitting transcriptomic analysis to study aging[29].

Spatially resolved sampling reveals information on the whereabouts of the biomolecules. With narrow focusing and adjustable power (fluence), lasers naturally became a method of choice in this research field. As illustrated in **Figure 1.4**, laser capture microdissection (LCM) concentrates infrared or ultraviolet light to lift off small tissue areas to groups of cells to single cells (e.g., 3–5 μm in size), with a capability for resolutions down to micrometers[28]. LCM-based HRMS facilitated biological studies on genes, proteins, and metabolites in broad contexts including cancer biopsies[45] and recently single-cell proteomics[46]. For smaller cell populations or dispersed cells, alternative strategies enhance the fidelity of sampling. Flow cytometry-based techniques, such as fluorescence-activated cell sorting (FACS), were employed in several studies to purify heterogeneous cell populations[47-49]. As illustrated in **Figure 1.4**, FACS enables accurate and high-throughput purification of cell populations to single cells based on fluorescent markers. Although specialized cytometers are able to sort particles ranging from 0.2 μm–150 μm in diameter, commercial FACS instruments usually have a low-pass limit of ~20 μm cell diameter[50].

ii. Subcellular Organization

For targeted subcellular analysis, distinct organelles may be isolated using various approaches. Sequential lysis of cellular compartments and density-based centrifugation are well-established for cells and cell types from multiple biological models. Organelles such as nuclei, mitochondria, peroxisomes, lysosomes, and the Golgi apparatus can be separated based on their size, shape, and surface charges[51]. Biochemical approaches and bioinformatics provide alternatives for subcellular analyses. Isotope tagging (LOPIT) and hyperLOPIT precluded the need for absolute organelle purification by taking a multipronged approach, in which organelles were partially separated into multiple fractions using multistep density centrifugation for multiplexed quantification via LC-HRMS, followed by deep machine learning to extract information on the subcellular localization of detected proteins[52]. MSI techniques such as micro- and nano-probe HRMS with bioinformatics emerged for precision organelle analysis, precluding the need for physically isolating microscopic organelles. Spatial single nuclear metabolomics (SEAM) integrated TOF-SIMS with novel computational data analysis modules (**Fig. 1.4**) to identify and localize single nuclei with variable metabolic patterns in healthy and fibrotic liver tissue[35].

iii. In Situ/Vivo

In situ sampling can take a molecular snapshot of the tissues or cells under innate conditions with minimal external perturbations during collection. As illustrated in **Figure 1.4**, the fast photochemical oxidation of proteins (FPOP) method utilizes laser-induced photolysis of hydrogen peroxide to irreversibly label proteins in their native state within a cell, thus enabling studies on protein structure and folding and protein-protein interactions[32]. Chemical labeling enables the isolation of molecular partners in interactions. Bioorthogonal reactions, for example, use a wide array of click chemistry to anchor specific biomolecules without disturbing the intracellular native environment. Bioorthogonal noncanonical amino acid tagging (BONCAT),

Figure 1.4, enabled the tracing of newly synthesized proteins by incorporating biotin-tagged noncanonical amino acids in the nascent proteins with follow-up affinity purification or imaging by fluorescence microscopy[53]. The approach recently allowed for tracing tumor cell proteome in ductal adenocarcinoma in the mouse[36].

Sampling *in vivo* extends HRMS 'omics to assess the molecular dynamics underlying physiological states. **Figure 1.4** shows examples of *in vivo* probes capable of minimally invasive sampling. Microdialysis was used to access molecular information on physiologically relevant timescales (e.g., from minutes to hours) in functional, behavioral, and clinical studies. Representative examples helped decipher the roles and spatiotemporal dynamics of neuropeptides and their role in feeding using crabs[54-55], decision-making and drug response in rats[56], and the tumor-promoting impact of radiation in glioblastoma using human xenografts in mice[57], as well as clinically for cerebral metabolic analysis of human patients with aneurysmal subarachnoid hemorrhage among others[58]. In the clinical setting, desorption electrospray ionization (DESI) and MasSpec Pen (**Fig. 1.4**) gained recognition for their ability to downscale microextraction into < 10 μL of water droplet on tissue regions of interest, even in morphologically complex samples[31, 59].

Microsampling confined to capillaries and nanochannels downscales sampling to the realm of cells and organelles. Whole cells, portions of a cell, or subcellular contents may be aspirated using precision-fabricated capillaries, usually with assistance from translation stages and microscopy. Early on, for single cells, live-video MS with direct ESI detected metabolites from single adherent embryonic fibroblasts (mouse) and plant cells[60-61], T-probe (**Fig. 1.4**) profiled metabolic perturbations in HeLa cells on anticancer treatment[62], and laser ablation electrospray ionization (LAESI) surveyed secondary metabolites in *Allium cepa* cells under guidance by

optical microscopy[63]. To detect deeper into single-cell 'omes, microsampling can be combined with high-efficiency separations prior to *in situ/vivo* HRMS. By aspirating <5% of the total cell volume using a pulled borosilicate capillary, proteomic and metabolomic cell heterogeneity was found in embryos of *X. laevis* and zebrafish[33, 64-66].

B. Sample Preparation

Sample preparation can make or break studies requiring trace detection, and those addressing spatiotemporal processes are no exception. Specimens encompassing few cells, single cells, or subcellular organelles provide only micro- to picogram quantities of starting materials[67-68]. Processing such limited amounts poses nontrivial challenges in handling minuscule volumes without solvent evaporation and with minimal-to-no losses for proteins, peptides, and metabolites due to surface adsorption[69]. The following section highlights developments that ushered HRMS to spatiotemporal 'omics from small tissue biopsies down to the single-cell and subcellular level with a special focus on proteomics.

i. Proteomics

HRMS provides critical information on proteomes during states of health and disease. Methods and technologies enabling bottom-up, middle-down, and top-down proteomics are able to identify and quantify broad types of proteins, assess their structure, gauge their interactions, and determine modifications during biological processes in space and time. Deep proteome coverage is an enabling feature of bottom-up proteomics, which is the most widely used of the three workflows. **Figure 1.5** shows the detailed workflow of bottom-up proteomics. In this approach, proteins are enzymatically digested into peptides, which can be efficiently ionized and sequenced to pinpoint their source proteins using advanced bioinformatics. Thousands of different proteins

can be identified at once using this approach, even from limited sample amounts. Complex spatiotemporal processes were studied using this approach at the organismal to subcellular levels in development[70-71], neuroscience[72-73], and various disease states[74-75]. The information from these measurements provides the foundation for hypothesis-driven and functional studies.

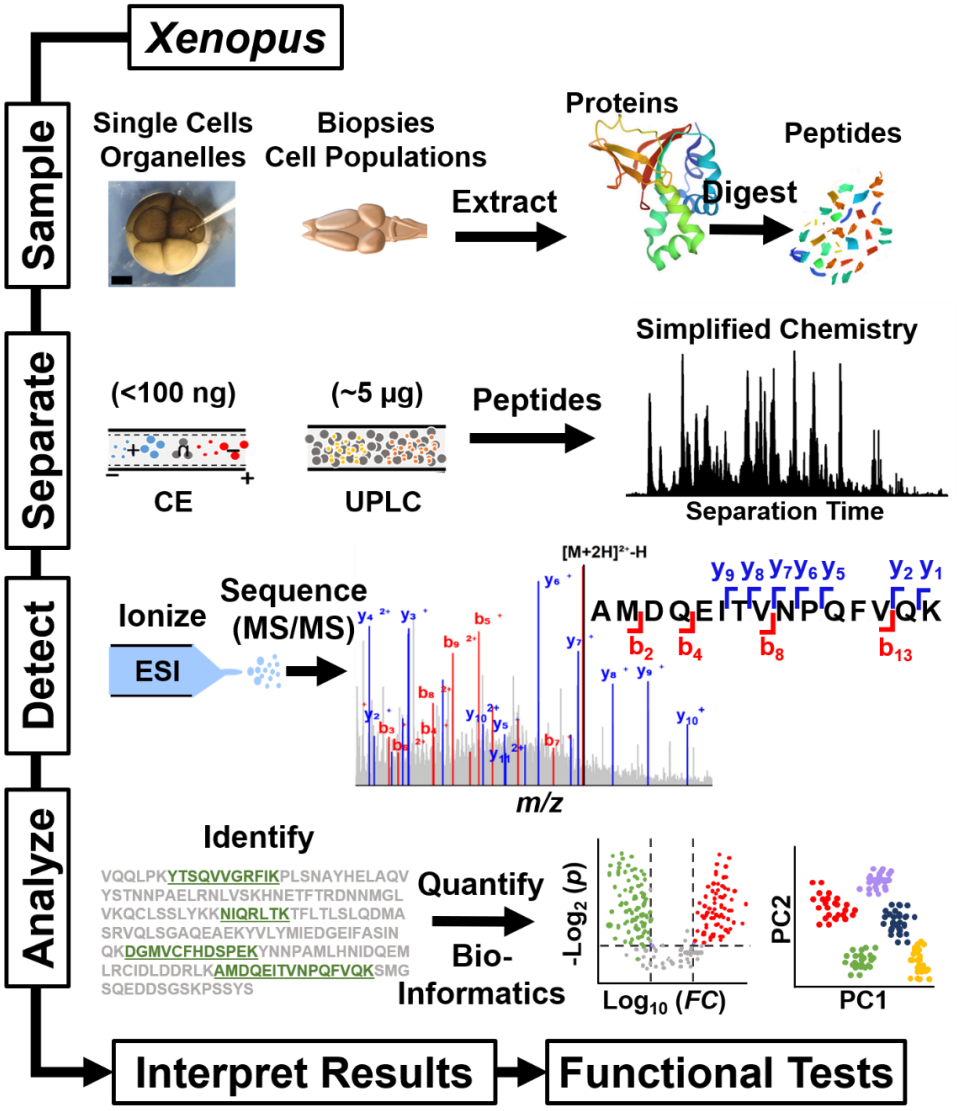

Figure 1.5. General steps of HRMS-based proteomics in *X. laevis* embryos and tissues. Proteins are identified and quantified based on the detection of proteotypic peptides. The sequenced peptide demonstrated is from protein Cop9. Discovery and targeted approaches enable high-throughput screening and biomarker discovery. Key: FC, fold change; p, statistical p-value (Student's t-test shown). Figures adapted with permission from[71].

HRMS-based proteomics executes multiple analytical steps with each able to affect sensitivity. Lysis, extraction, reduction, alkylation (of disulfide bonds), enzymatic digestion, and desalting are routine stages of sample processing requiring care for limited samples. For single-cell measurements, for instance, MS-compatible detergents and mechanical lysis methods (i.e., freezing/thawing or laser pulses) and exclusion of reduction and alkylation eliminated the need for downstream clean-up, reduced nonspecific adsorptive losses, and dilution[33, 76]. Select strategies enabling high-fidelity proteomics are presented in **Figure 1.6** and discussed further.

"One-pot" strategies confined the analytical steps to a single location. For amount-limited samples, immobilization of molecular processing and reduction or elimination of sample transfers proved particularly beneficial. For example, gel-aided sample preparation (GASP)[77], suspension trapping filter (S-Trap)[78], and filter-aided sample preparation (FASP)[79] used gels and filter membranes, respectively, to reduce sample transfer steps and improve coverage from hundreds of micrograms of protein. By miniaturizing and confining sample processing using a multi-well plate, micro-FASP was able to identify ~3,400 proteins from ~200 ng of the yolk-free proteome from single blastomeres that were isolated from early-stage *X. laevis* embryos[80]. As illustrated in **Figure 1.6**, single-cell proteomics by MS (SCoPE-MS) limited sample processing to one pot, allowing for multiplexed protein quantification among single cells[81].

Capillary-based systems may serve as alternative substrates for single-cell sampling and processing. As shown in **Figure 1.6**, The integrated proteome analysis device (iPAD) directly supplied a suspension of living cells (iPAD-100) for online cell lysis, protein digestion, and peptide desalting and trapping to a 10-port valve-based system on a commercial nanoLC, identifying ~600-650 proteins from a single injection of 100 cells (~10 ng protein) using nanoLC-HRMS[82]. In pursuit of cell heterogeneity, iPAD-1 (**Fig. 1.6**) used an ultranarrow-bore

analytical column to identify ~270 proteins from a single HeLa cell (220 pg proteome) in 1 h of analysis with nanoLC-HRMS[83]. Microfabricated devices present a natural setting for "one-pot" workflows toward sensitivity and advanced measurement throughput. As illustrated in **Figure 1.6**, nanodroplet processing in one pot for trace samples (nanoPOTS) miniaturized the processing of 100 cells to a single mammalian cell onto nano-wells on microfabricated glass chips using nanoliters of reagents for deep proteomics[84-85]. NanoPOTS also enabled proteomics on single sensory hair cells measuring ~0.6–1 pL in volume, facilitating the study of cell differentiation in chicken embryos[86]. Nested nanoPOTS (N2) arrayed multiple wells to enhance measurement throughout for multiplexed quantification using nanoLC-HRMS[87]. Recent efforts sought to further refine one-pot sample preparation strategies to make the setups cost effective and allow scalable sample preparation, as was recently reviewed in[88]. Automated preparation in one pot for trace samples (autoPOTS), the latest upgrade of nanoPOTS, now allows automated processing of 1–500 cells using only commercially available instrumentation for sample processing, making it widely adaptable[89]. Similar strategies also emerged for top-down proteomics. Detection of intact proteins provides insights into proteoforms and PTMs[90]. Freeze-thaw lysis and non-ionic surfactant-based extraction from ~200 brain cells enabled spatial top-down proteomics for ~400 proteoforms between the optic tectum and telencephalon regions of the zebrafish brain[91]. Further advancements in technology raise exciting potentials to leverage proteoforms as biomarkers of diseases[92].

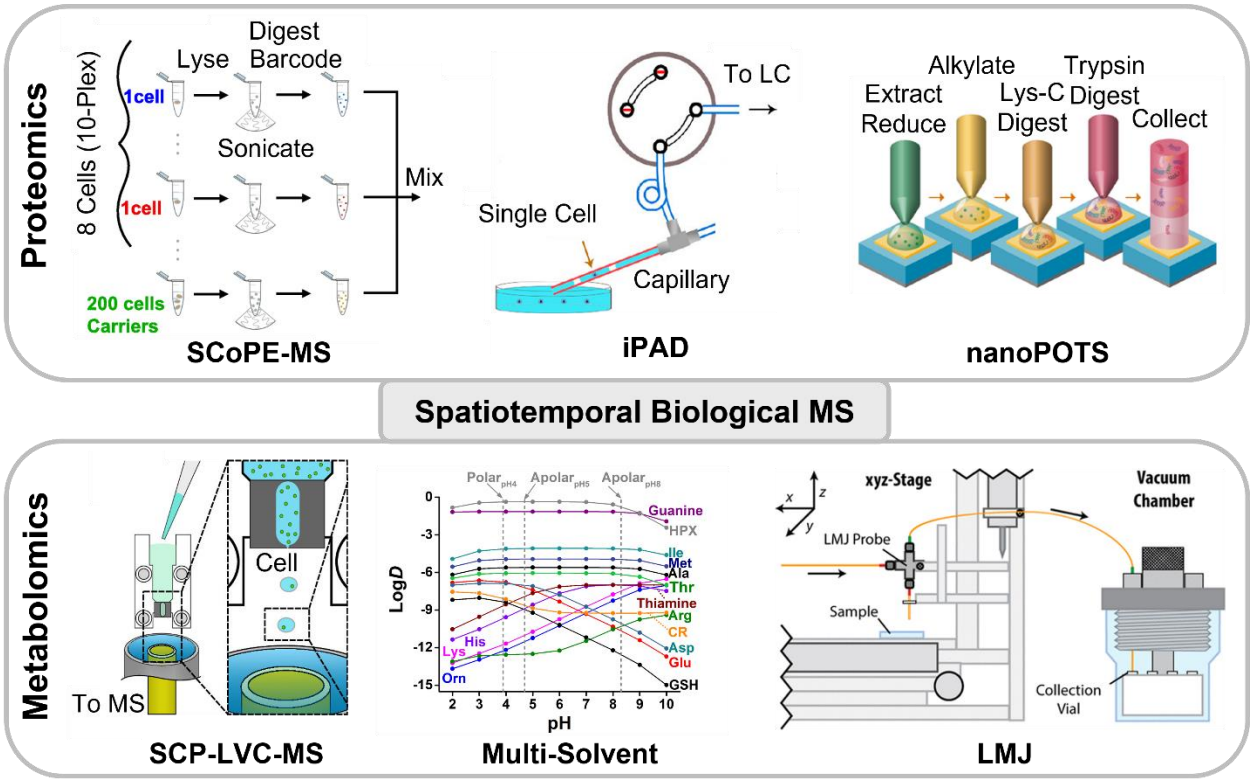

Figure 1.6. Scalable and high-sensitive sample processing for metabolites and proteins enabling spatiotemporal biological MS. One-pot sample processing (SCoPE), direct cell capture and processing (iPAD), and miniaturization of sample processing steps (nanoPOTS) deepen the detectable proteome by reducing analyte losses. Direct cell capture (SCP-LVC), multi-solvent processing, and in situ, metabolite extraction (LMJ) after MALDI imaging improves coverage of the single-cell metabolome. Technological abbreviations are in the text. Adapted with permission from[80-81, 93-96].

ii. Peptidomics and Metabolomics

While analyzing enzymatically digested peptides (i.e., bottom-up proteomics) provides information about the proteomic content of a cell or biological sample, endogenous peptides, such as neuropeptides, can provide additional insight into key biological processes, particularly related to the nervous and neuroendocrine systems. These peptides are synthesized from larger proteins via enzymatic cleavage, and their biological activity is often distinct from the preprohormone or precursor-related peptides[97]. Peptidomics consists of peptide extraction, filtration, via molecular weight cutoff filtration, and desalting prior to MS analysis. A critical

aspect of this process is the deactivation of proteolytic enzymes to prevent peptide degradation and achieve highly sensitive peptidomics analyses, which can be accomplished by flash-freezing the tissue immediately after dissection, heat stabilization, or acidification. Acidification is particularly effective for deactivating proteases when the acid is in an organic solvent (e.g., acidified methanol or acidified acetonitrile), as while proteases are deactivated by the low pH, large proteins are precipitated from the solution. For example, acidified methanol was recently used to successfully extract neuropeptides from single mollusk neurons, enabling sequencing of peptides in the cytoplasm and nucleus of the neurons and identification of differences in peptide abundances based on subcellular location[98]. Alternatively, MSI strategies are popular for direct measurement and spatial localization of bioactive peptides, particularly in limited samples; sampling with a focused laser beam requires minimal sample preparation yet yields a sufficient amount of analytes for ionization and detection. Development of validated protocols made sample preparation for MSI straightforward and robust[99-101].

High-sensitivity MS-based metabolomics can be streamlined to deliver high reproducibility facilitating both absolute and relative quantification of metabolites. Common sample preparation protocols in metabolomics include quenching of enzymatic activity, analyte extraction, purification, desalting, and/or chemical derivatization prior to separation. To so deepen metabolite coverage, the detection sensitivity can be improved by tailoring the chemical composition, pH, and temperature of the extraction solvent to match the chemical properties of the compound classes of interest (e.g., amino acids, lipids, nucleotides). For example, a discovery study in *E. coli* culture reported that polar solvents (e.g., water, ethanol) in combination with high temperature effectively extracted hydrophobic and hydrophilic metabolites as compared to less polar solvents (e.g., acetone)[102]. Several advances for cellular

and subcellular sample preparation were developed, including optimized extraction solvent mixtures[93], on-column preconcentration or sample stacking[103-105], and microfluidic devices[106].

C. Separation

To facilitate molecular detection in the collected specimens, chemical separation usually precedes HRMS measurement. Separation provides another layer of compound-dependent information, the time of separation, to aid molecular identifications. It also helps minimize or eliminate chemical interferences during ionization and detection, thus improving sensitivity and the confidence of molecular identifications. Further, by spreading the chemical complexity of the sample over time, separation enhances the efficiency of duty cycle utilization during tandem and multi-stage MS, which in turn deepens the detectable and quantifiable portion of the 'omes. This section highlights approaches, shown in **Figure 1.7**, that facilitate spatiotemporal studies on limited specimens, ranging from cell populations to single cells to organelles, using some form of separation.

i. Liquid Chromatography

As a result of extensive research in stationary phase chemistries and engineered systems capable of withstanding high pressures, automation, and commercialization, LC became the mainstream separation technology for HRMS. Virtually all types of proteins, peptides, metabolites, lipids, and drug molecules can be separated using high-performance LC, or HPLC[107-108]. Reversed-phase liquid chromatography (RPLC) employing packed-bed stationary phases with C_{18} chemistries (**Fig. 1.7**) is efficient in resolving peptides and hydrophobic metabolites including lipids, whereas C_4 or C_8 chemistries were found to work well for intact proteins[90]. The

development of hydrophilic-interaction LC (HILIC) extended analyses to small and polar molecules, including neuropeptides, amino acids, and glycans[107, 109].

Combining orthogonal mechanisms of separation enhances trace-sensitive analyses. Higher theoretical plate numbers resulting from the hyphenation of multiple dimensions of separation deepen the molecular depth of detection. The multidimensional protein identification technology (MudPIT) recognized these fundamental advantages for bottom-up proteomics early on by supplementing strong cation exchange (SCX) with reversed-phase chromatography[110]. To reduce the salt load from SCX, a modification of the approach used RPLC chemistries in both dimensions, first to fractionate at high pH (e.g., pH 8–10), then to separate each fraction at the conventional low pH (e.g., pH 2–3). High pH fractionation doubled the overall depth of the detectable phosphoproteome from mouse cell lines compared to SCX, identifying over 30,000 phosphopeptides[111]. For limited cell populations, for example, this approach recently yielded a 70% improvement in the coverage of the measurable proteome, allowing for the molecular characterization of the developing neural ectoderm during embryogenesis[42].

Recent advances in high performance LC (HPLC) extended HRMS to ever-decreasing amounts of materials (**Fig. 1.7**). To reduce analytical requirements to a few micrograms to nanograms of proteomes, new-generational columns enhanced the peak capacity of partition chromatography by packing narrow-bore columns (inner diameters < 50 µm) with small silica particles (< 5 µm). These columns operate at nanoflow rates (< 50 nL/min), thus concurrently reducing sample dilution and increasing ionization efficiency towards ultrasensitive detection. Recently, 362–874 protein groups became identifiable from a single HeLa cell (< 150 pg) using flow rates of ~20 nL/min through an in-house-packed nanoLC column with 20 µm inner diameter[112]. With 10–100-fold improvement in sensitivity over narrow-bore packed columns (~15–30 µm inner

diameter), these continuous developments in HPLC were recently adopted to microfluidic systems capable of automating lossless sample preparation, such as nanoPOTS[113], to address the challenge of protein loss.

New-generation LC columns overcame challenges associated with packing narrow columns with robust performance. As illustrated in **Figure 1.7**, open tubular, monolithic, and micropillar array columns (µPAC) reduced band broadening and increased peak capacities using flow rates decreased to nanoliters to sub nanoliters per minute[114]. The resulting enhancements in separation improved sensitivity and reproducibility for trace-level proteomics and peptidomics[107-108, 115] as well as metabolomics[116-118]. Capillary monolithic columns with a 12 nL/min flow rate enabled the identification of over 2,300 proteins from 1 ng of HeLa digest[119]. Flows on the order of picoliters per minute through narrow open tubular columns with ~2 µm inner diameter recently elevated identifications to ~1,000 proteins from 75 pg of *Shewanella oneidensis* tryptic digest, which was equivalent to half the amount obtained from a single HeLa cell[120].

Improvements in LC-ESI interfaces played an equally important role in advancing the sensitivity of analyses. In the conventional ESI interfaces, the LC eluent is fed through a metal emitter or pulled glass capillary, and sufficient potential difference is established between the spray tip and the mass spectrometer to form a stable electrospray. **Figure 1.7** shows the impact the onset of different spraying regimes has on spray dynamic and MS ion signal when the ESI potential was adjusted. The relative ion signals, stability, and chemical composition of the electrospray are all impacted by alterations in the spraying regime. As illustrated, the cone-jet regime generates small droplets continuously, thus maximizing ionization efficiency[121]. The interface called UniSpray (**Fig. 1.6**) places an impactor rod of high voltage between the grounded capillary tip and inlet of the MS to facilitate electrospray ionization of the primary electrosprayed droplets,

thus boosting S/N for neuropeptides[122]. Following the scaling laws of ESI, emitter tips with submicron-diameters (e.g., ~0.6 μm) were recently demonstrated to improve protein and oligonucleotide sensitivity by efficiently reducing their formation of clusters and adducts[123].

ii. Capillary Electrophoresis

While LC is a powerful separation method for complex biological samples, separation with CE can offer distinct advantages. Unlike HPLC, CE separates charged compounds in solution based on differences in their electrophoretic mobilities. With typical sample consumption on the order of a few nanoliters, this technology is a natural fit for volume- and mass-limited samples. Several methods allow for efficient pre-concentration of analytes via in-column sample stacking[124]. With large-volume sample stacking (LVSS) using a dynamic pH junction, a 10-fold enhancement in sensitivity (**Fig. 1.7**) enabled the detection of 5,700 proteoforms in the *Escherichia coli* proteome[125] and peptides from the renin-angiotensin system in the paraventricular nucleus of the mouse central nervous system[126]. Ultrasensitive measurements also benefited from the innovation of specialized CE ion sources for HRMS. The development of low sheath-flow and sheathless interfaces supported the coupling of CE with electrospray ionization (ESI) HRMS, reviewed elsewhere[127-128]. **Figure 1.7** highlights representative CE-ESI interfaces that enabled stable and sensitive CE-ESI operation using blunt-tip and tapered-tip metal as well as microfabricated borosilicate and fused silica emitters. These custom-built CE-HRMS instruments extended these measurements to attomole to zeptomole to yoctomole levels of metabolites and peptides from limited samples and broad types of biological models[126, 129-134]. Integration of CE with orthogonal HRMS technologies opened exciting new directions in basic research. CE-HRMS with CE modes[135] and LC fractionation[136] deepened the detection of proteomes.

iii. Ion Mobility Spectrometry

Gas-phase ion separation provides speed and added resolution to aid molecular detection. With ion mobility separations taking less than 100 milliseconds, ion mobility can be nested into most liquid-phase separations prior to HRMS detection. Most IMS techniques yield information to enable the calculation of collision cross sections (CCS), which serves as a compound-dependent information to aid molecular identifications. This approach enabled the elucidation of compound classes with large molecular diversity, including D- and L-amino acid containing peptides[137] and neuropeptides of the renin-angiotensin system in the paraventricular nucleus of the mouse central nervous system[126]. IMS techniques combined with other separation techniques, such as HPLC, to help distinguish isomers, isobars, and conformers. As a dual separation device and charge-state filter, field asymmetric waveform IMS (FAIMS) increased unique peptide identifications by ~30% and doubled the number of quantifiable peptides over non-IMS experiment from HEK293 cells due to a 10-fold enhancement in ion signal abundance and improved utilization of MS/MS duty cycles[138].

D. Mass Analysis and Detection

Spatial HRMS ‘omics profits from a still-expanding array of technological innovations. Development of specialized ion optics enabling efficient ion collection and transfer (e.g., ion funnel) as well as systems integrating new-generation mass analyzers, digitizers, and ion detectors (e.g., TOF, orbitrap, FTICR, SIMS) ushered in a new era of sensitivity (detection limit), speed, and spectral resolution for limited specimens. Spatiotemporal ‘omics naturally benefited from advances in tandem MS with increasing throughput, sensitivity, and specificity. An intelligent data dependent acquisition (DDA) method monitored the elution order of peptides

based on prior discovery measurements to target tandem MS events, resulting in 80% more protein identifications than the standard[139]. In parallel, instruments with multiple fragmentation cells opened the door to using structurally diagnostic fragment ions to trigger sequential or multiplexed dissociation toward deeper sequence coverage. The advent of data independent acquisition (DIA) remedied limited tandem MS duty cycles during DDA, approximately doubling the number of neuropeptides in limited neural tissues[140].

Spatial HRMS provides several strategies for quantification (recall **Fig. 1.3**). Both label-free quantification (LFQ) and label-based quantification were adapted to limited specimens in the context of spatiotemporal biology. LFQ uses spectral counts, peak intensities, or peak areas to estimate the abundance of proteins, peptides, and metabolites without barcoding[141-142]. A reduction in sample processing steps aids analytics on trace amounts of samples, as was reviewed recently[143]. To enhance throughput, label-based quantification may integrate stable isotopes during sample preparation *in vivo* (metabolic labeling) or *in vitro* (chemical labeling) to enable multiplexing, as was reviewed elsewhere[144]. Label-based quantification became quasi routine for high-throughput quantification across broad spatial and temporal processes, lately even for single cells. The carrier approach pioneered the use of tandem mass tags (Thermo Fisher Scientific) in abundance to boost tandem MS on isobarically tagged protein digests from single cells[88], while the usage of a sacrificial protein helped reduce nonspecific protein losses during sample preparation[143].

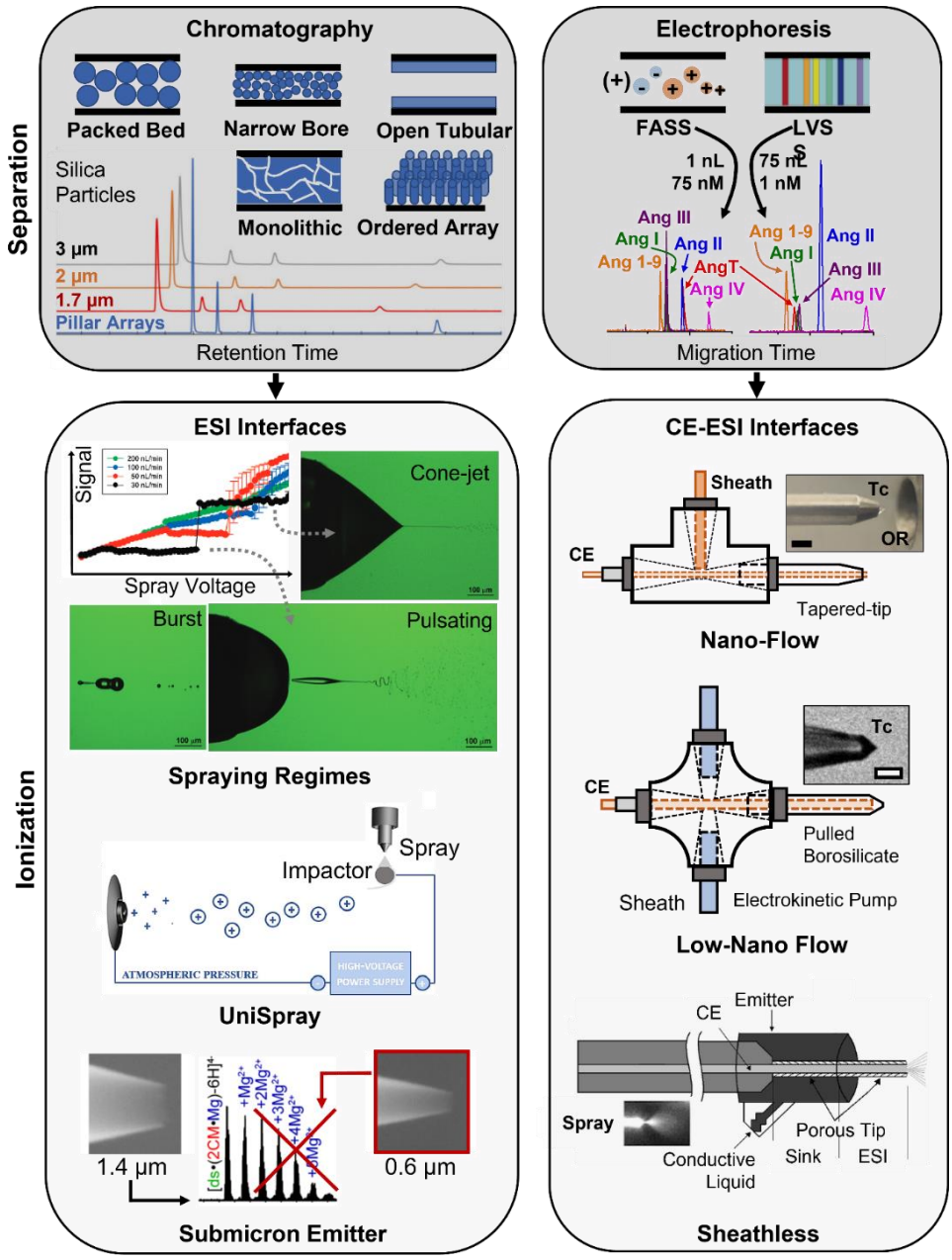

Figure 1.7. Advances in separation and ionization aided the chemical analysis of limited amounts of materials. Reducing the size of packing material, the diameter of LC columns, and newer generation columns minimize the dispersion of molecules for high sensitivity during chromatography. Diverse stacking methods boost sensitivity during capillary electrophoresis. Specialized technologies and methods improve ion yields during electrospray ionization by controlling the spraying regime, charging of droplets, and reducing analyte dilution. Technological abbreviations are in the text. Key: OR, orifice of mass spectrometer; Tc, Taylor cone; FASS, LVS abbreviations are in the text. Adapted with permission from[33, 114, 121-123, 126, 129, 131, 145-146].

E. Data Analysis

Bioinformatics plays a critical and enabling role in spatiotemporal HRMS. This field of HRMS yields challenging data, not only in terms of complexity and size but also metadata to assess in relation to space and time. General steps for data analysis include raw data processing (feature extraction or identification), data preprocessing, statistical analysis (quantitative information), and gene ontology (GO) annotations (qualitative information). Data from imaging microscopy (optical to electron) and functional tests, including electrophysiology and behavioral assays, complement chemical data from MS ‘omics with information on the biological phenotype. Chemometrics is pivotal for the interpretation of metadata from spatiotemporal HRMS (recall **Fig. 1.3**). This field of research offers various solutions for statistical and multivariate analyses. For example, volcano plots are widely used to examine molecular differences between spatially or temporally distinct sample types. Unsupervised and supervised dimensionality reduction and clustering algorithms such as principal component analysis (PCA), linear discriminant analysis (LDA), t-distributed stochastic neighbor embedding (t-SNE), hierarchical cluster analysis (HCA), heatmaps, k- and fuzzy-mean clustering are broadly adaptable to exploring systematic molecular differences between samples, such as types of tissues or cells. Software packages like MetaboAnalyst[147], Perseus[148] (Tyanova et al., 2016), Orange[149] and IOAT[150] provide user-friendly platforms for performing such analyses. These and other statistical and data visualization pipelines help generate and test new hypotheses at the levels of proteomes, peptidomes, and metabolomes in biological processes occurring in space and time.

F. Summary

Interdisciplinary efforts molded HRMS ‘omics into a powerful tool to advance basic and translational research and also understanding of biological processes occurring in space and time. These developments are the result of still rapidly progressing interdisciplinary and collaborative efforts innovating essentially all steps of biological HRMS. Inventions of specialized strategies, equipment, and miniaturization to collect and process samples with speed and high fidelity is pivotal for studies on limited and dynamic biological systems. Advances in chemical separations, ionization, ion transfer and accumulation, and HRMS detection increase detection sensitivity and measurement throughput, thus ushering ‘omics to increasingly finer spatial and temporal resolutions. Continued developments in data acquisition and user-friendly software packages to handle the resulting complex data promise to further deepen the detectable and quantifiable portion of proteomes, peptidomes, and metabolomes from trace amounts of specimens. Grand challenges in the acquisition, integration, and distribution of complex data from spatiotemporal multi-‘omics pose grant challenges and exciting opportunities to move science forward toward a holistic understanding of life processes.

Chapter 2: Dilute to Enrich for Deeper Proteomics: A Yolk-Depleted Carrier for Limited Populations of Embryonic (Frog) Cells

(This chapter is reproduced and reused with permission from Pade, L. R., Lombard-Banek, C., Li, J., & Nemes, P. (2024). Dilute to Enrich for Deeper Proteomics: A Yolk-Depleted Carrier for Limited Populations of Embryonic (Frog) Cells. Journal of Proteome Research, 23(2), 692-703.)

Contribution: L.R.P. collected and prepared the samples for proteomic analysis, analyzed the samples via LC-MS, performed data analysis, interpreted the results, drafted the manuscript, and revised the contents.

2.1 Abstract

Abundant proteins challenge deep mass spectrometry (MS) analysis of the proteome. Yolk, the source of food in many developing vertebrate embryos, complicates chemical separation and interferes with detection. We report here a strategy that enhances bottom-up proteomics in yolk-laden specimens by diluting the interferences using a yolk-depleted carrier (YODEC) proteome via isobaric multiplexing quantification. This method was tested on embryos of the South African Clawed Frog (*Xenopus laevis*), where a >90% yolk proteome content challenges deep proteomics. As a proof of concept, we isolated neural and epidermal fated cell clones from the embryo by dissection or fluorescence-activated cell sorting. Compared with the standard multiplexing carrier approach, YODEC more than doubled the detectable *X. laevis* proteome, identifying 5,218 proteins from D11 cell clones dissected from the embryo. Ca. ∼80% of the proteins were quantified without dropouts in any of the analytical channels. YODEC with high-pH fractionation quantified 3,133 proteins from ∼8,000 V11 cells that were sorted from ca. 2 embryos (1.5 μg total, or 150 ng yolk-free proteome), marking a 15-fold improvement in proteome coverage vs the standard proteomics approach. About 60% of these proteins were only quantifiable by YODEC, including molecular adaptors, transporters, translation, and transcription factors. While this study was tailored to limited populations of *Xenopus* cells, we

anticipate the approach of "dilute to enrich" using a depleted carrier proteome to be adaptable to other biological models in which abundant proteins challenge deep MS proteomics.

2.2 Introduction

Abundant proteins are a major challenge to the deep characterization of the proteome. In MS, the modern technology of choice for proteomics, high-concentration proteins interfere with essentially all steps of the analytical workflow. During the bottom-up, most common approach, peptides that result from the digestion of abundant proteins increase the chemical complexity of the sample, limit the efficacy of molecular separation, and suppress the ionization of low-abundance signals.[151] Concurrently, the resulting increase in the dynamic concentration range weakens the collectable signal due to dilution and promotes the exclusion of low-intensity ions from sequencing in favor of those of high concentration during data dependent acquisition (DDA). Abundant ion signals can also be redundantly fragmented, further burdening the duty cycle of $MS{-}MS^n$ scans. Several strategies have been developed to remedy these challenges. For example, multidimensional separation seeks to spread sample complexity over time to better utilize the bandwidth of tandem MS[152]. Specialized data acquisition methods can also help reduce detection biases through the use of exclusion lists,[3] a nested DDA ladder,[4] advanced precursor ion selection,[5] DIA acquisition,[6] or wide-window DDA.[7]

Challenges in protein detection due to abundant molecules transcend across the taxa. Ca. 75% of the mammalian proteome is comprised of ~2,300 housekeeping proteins of high abundance, including but not limited to albumins, immunoglobulins, haptoglobins, and transferrin in tissues and plasma.[153] In the popular embryonic models of cell and developmental biology and neuroscience, such as the frog *Xenopus* species[154], zebrafish[155], and chick[156], the interfering

proteome is generally molecularly less diverse, nonetheless more severe than their mammalian counterparts. For example, yolk proteins make up >90% of the total proteome by concentration at the early stages of development in *X. laevis.*[157-159] The measurable depth of proteome coverage consequently suffers in many ways, primarily driven by interferences during chemical separation, ionization, and detection.[159-160] Despite significant efforts in developing methods and technologies, interference from abundant proteins has remained a bottleneck for deep MS proteomics.

Protein depletion has long offered a successful means of mitigating such interferences. Various techniques have been introduced to remove abundant proteins for discovery-based or targeted MS, especially in serum and plasma. The method of Digestion and Depletion of Abundant Proteins (DigDeAPr) leveraged abundance-dependent digestion kinetics to reduce abundant proteins in plasma.[161-162] The popular dye Cibacron Blue binds to albumin, helping to remove ∼40−50% of this protein in serum or plasma. Nonetheless, nonspecific depletion has been reported, for example, causing losses to the signaling proteins such as junction plakoglobin and complement component C7 and C5.[163] Alternatively, perchloric acid is commonly used to precipitate plasma proteins with simplicity and high throughput[164]. To enhance the molecular specificity, immunoaffinity reagents may leverage antibody−antigen interactions for protein capture on a stationary phase.[165-166] Likewise, the M13 phage library has been conjugated to magnetic microspheres to enrich the plasma proteome.[167] Immunoaffinity liquid chromatography (LC) columns may integrate multiple antibodies to deplete the 7−14 most abundant proteins, thus improving LC-MS protein identifications by 25% relative to the crude human plasma.[168] Spin columns employing similar biochemistry provide a convenient platform for removing 14 targeted proteins as part of the modern proteomics workflow.[169] Offering non-specific fractionation,

columns integrating strong cation exchange and reversed-phase (RP) chemistries have also been introduced.[170]

Approaches based on the isolation of subcellular organelles can also simplify the proteome. Several membranous organelles and nonmembranous assemblies compartmentalize in eukaryotic cells. Such components can be fractionated or enriched using ultracentrifugation or gradient centrifugation.[171] For example, the isolation of nuclear, cytoplasmic, and cell membrane fractions has enabled spatial mapping of proteins in a steady state and activated human T cells.[172] Mitochondrial enrichment from rodent heart and muscle cells has permitted the characterization of the local mitochondrial proteome due to respiratory insufficiency.[173] Similarly, in the yolk-laden amphibian, fish, and chick models, analytical workflows may exploit the innate biology of protein storage to deplete abundant proteins. For example, zebrafish and chick embryos can be readily dissected away from the yolk sac. In the *Xenopus* species, yolk makes up >90% of the total proteome in every cell of the embryo to fuel early development, where it is segregated into platelets of 1−12 μm size, dispersed throughout each cell.[158] Physical methods, such as sucrose-based centrifugation, have enabled the purification of the yolk storage.[159, 174] This approach, commonly referred to as “yolk depletion,” has helped reduce yolk content to ∼30%, boosting identifications from 679 to 1,727 proteins in cell lineages in the developing *X. laevis* embryo.[14] However, organelle isolation, including yolk depletion, does not scale well to small amounts of specimens.

A fast-growing number of investigations seek to measure limited proteomes in high[175]-to-ultrahigh sensitivity, (partially reviewed in the literature[29−31]) for example, from tissue biopsies, small cell populations, single cells, and even subcellular domains. Consequently, various strategies have been developed to advance sample preparation and MS detection to increasingly

smaller amounts of starting materials. One-pot and microfluidic techniques such as nanodroplet processing in one pot for trace samples (NanoPOTS)[176-177] limit adsorptive losses, sample transfer, and processing volumes. Recently, microPOTS has identified over 1,500 proteins from 200 esophageal cells, delineating stress-induced proteomic changes.[178] Similarly, filter aided sample preparation (FASP) confined the steps of sample processing and cleanup onto a filter membrane, detecting ~4,000 membrane proteins from a single mouse hippocampus.[179] In the method Single Cell Proteomics (SCoPE),[180] multiplexing barcoding has been utilized with a "carrier" proteome channel to quantify ~3,000 proteins among ~1,400 single macrophages and monocytes.[181] Although these strategies have improved sample preparation and data acquisition toward ever-lower sample amounts, their designs do not address interferences from abundant proteins.

Our goal in this study was to deepen detection and quantification toward low-abundance proteins from limited cell populations in the *Xenopus* species, where a dominant subcellular yolk proteome has historically hampered MS proteomics. We selected *X. laevis* as the biological model due to its important implications for health studies (reviewed in the literature[9,38,39]), including but not limited to normal and diseased development[38] and emerging prominence in cell differentiation,[40] tissue induction,[41] regeneration,[42] and neuro- development.[43] Our previous depletion of yolk platelets reduced more than ~90% of yolk content to ~30% and improved proteome detection.[14] Through mathematical modeling, we predicted that spiking of the endogenous proteome with an isobarically tagged yolk-depleted (YD) proteome dilutes yolk peptides faster than nonyolk peptides. We reasoned that this approach, which we termed yolk-depleted carrier (YODEC), would therefore enhance detection and quantification for nonyolk and low-concentration proteins from limited populations of cells, where mechanical depletion

becomes too difficult to scale. We developed and validated the analytical framework of YODEC using the commercial HeLa proteome as the benchmark. The methodology was demonstrated on embryonic cells (blastomeres) that are contained in neural and epidermal tissue fated cell populations we dissected or isolated by fluorescent-activated cell sorting (FACS). Our results from this study suggest that the concept of "dilute to enrich" in YODEC can be generalized and applicable also in studies where limited amounts of specimens with interference from abundant proteins challenge the measurable proteome coverage.

2.3 Experimental Section

Materials and Reagents. LC-MS-grade solvents and reagent-grade chemicals were purchased from Fisher Scientific (Pittsburgh, PA) or MilliporeSigma (St. Louis, MO). Trypsin protease (MS- -grade), HeLa protein digest standard (Pierce), TMT10plex isobaric labeling reagent set, and green fluorescent dextran (FD, Alexa Fluor 488) were purchased from Fisher Scientific.

Solutions. For embryo cultures and dissections, 100% (w/v) and 50% Steinberg's solutions (SS) were prepared according to established protocols.[182] For FACS, embryos were dissociated in the Modified Newport 2.0 buffer[174] and suspended in 3−4 mL of 1× phosphate buffered saline (PBS) containing 2% (w/v) bovine serum albumin (BSA). The yolk depletion buffer was prepared following previous publications.[159, 183] The conventional proteome lysis buffer contained 1% (v/v) sodium dodecyl sulfate, 150 mM NaCl, 5 mM EDTA, 20 mM Tris hydrochloride, and a protease inhibitor cocktail (Halt, Thermo) at 1× final concentration in the buffer.

Animal Care and Embryology. A breeding colony of sexually mature adult male and female *X. laevis* frogs (Nasco, Fort Atkinson, WI) was maintained and handled humanely following protocols approved by the Institutional Animal Care and Use Committee of the University of

Maryland (approvals R-DEC-17-57 and R-FEB- 21-07). Embryos were obtained via gonadotropin-induced natural mating and dejellied as reported elsewhere.[182] Two-cell embryos with stereotypical pigmentation were chosen and cultured to the 16-cell stage (Nieuwkoop and Faber stage 5)[184] in 100% SS. The neural-tissue-fated animal-dorsal midline (called D11) and epidermally destined animal-ventral midline (called V11) cells were identified based on pigmentation, size, and location in the embryo in reference to established and reproducible cell fate maps in this model.[8] Their descendent lineages were fluorescently traced via injection of the Alexa Fluor 488 dye following our previous work.[185] The embryos were cultured in 50% SS and monitored under a stereomicroscope using epifluorescence (model SMZ18 with a GFP- B 480 filter, Nikon Instruments Inc., Melville, NY).

Cell Collection and Proteome Preparation. We tailored the methods of cell isolation to the goal of each experiment in this study. YODEC was developed and refined using the D11 cell clone, which we were able to dissect reproducibly from stage-13 (neurula) embryos under a fluorescence stereomicroscope. The YODEC approach was scaled to the fluorescently labeled V11 lineage that required isolation using FACS, yielding ∼20% starting proteome material compared to the D11 tissue biopsy. For FACS, the embryos were dissociated in the Modified Newport 2.0 buffer and the resulting cell suspension was transferred to LoBind tubes (Eppendorf) for pelleting at 400*g* at 4 °C for 5 min, before suspension in 1× PBS containing 2% BSA (to prevent aggregation or cell adsorption on vial surfaces). The fluorescently labeled cells were sorted into glass vials containing 1× PBS with 2% BSA. The cells were pelleted (400*g* at 4 °C for 5 min) and rinsed with 1× PBS three times to remove excess BSA. The remaining PBS solution was removed by pipetting, and the resulting cells were pelleted before storage at −80 °C until further processing. Whole embryos at the same developmental stage (stage 13) were used to

prepare the “carrier” proteome. The yolk platelets from the whole embryo were isolated by centrifugation in a sucrose buffer following previous protocols.[159, 183] The dissected tissues, sorted cells, and carrier proteome were processed conventionally for bottom-up proteomics. The peptide mixtures resulting from trypsin digestion were barcoded with TMTs following the experimental conditions provided in **Table 2.1**. Optionally, the chemical complexity of the sample was reduced via high-pH RP fractionation in a spin column following a protocol established by the vendor (Pierce kit no. 84868, Thermo Fisher Scientific).

Table 2.1. Design of multiplexing barcoding experiments using tandem mass tags (TMTs, Fisher).

Isobaric barcoding for the HeLa proteome digest and D11 lineage dissected from X. laevis (stage 13):

Experiment	***X. laevis* D11 or HeLa Analyte Proteomes**				**Carrier/Analyte Proteome Ratio**
TMT channel	126	127N	128C	129N	130C
No carrier	1.5	1	1	1.5	No
1× carrier	1.5	1	1	1.5	1×
2× carrier	1.5	1	1	1.5	2×
4× carrier	1.5	1	1	1.5	4×

Isobaric barcoding for the V11 lineage that was sorted from X. laevis (stage 13) by FACS:

Experiment	***X. laevis* V11 Proteome**			**Carrier/Analyte Proteome Ratio**
TMT channel	126	127N	128C	130C
No carrier	1	1.5	1	No
Carrier: No Fractionation	1	1.5	1	4×
Carrier: With Fractionation	1	1.5	1	4×

Mass Spectrometry. About 1 μg of total TMT-tagged peptides was separated on a nanoflow liquid chromatograph (Dionex Ultimate 3000 RSLC, Thermo). The peptides were loaded onto a C18 trapping column (100 μm inner diameter, 5 μm particle size with 100 Å pores, 2 cm length, Acclaim PepMap 100, Thermo) in Buffer A (LC-MS grade water, 0.1% formic acid) at a rate of 10 μL/min for 10 min. RP separation was conducted on a C18 column (75 μm inner diameter, 3 μm particle size with 100 Å pores, 25 cm length, Acclaim PepMap 100, Thermo) using a 150 min stepped gradient ramping Buffer B (LC-MS grade ACN, 0.1% formic acid) at 300 nL/min as follows: 5% for 0−5 min, to 10% in 10 min, to 26.5% in 135 min, to 80% in 2 min, held at 80% for 3 min, then ramped to 5% in 5 min, and equilibrated at 5% for 15 min.

The peptides were electrospray-ionized and detected on a quadrupole-ion trap-Orbitrap tribrid ultrahigh-resolution mass spectrometer (Orbitrap Fusion Lumos, Thermo). The nano- flow electrospray ionization (ESI) source was prepared from a tapered-tip stainless-steel emitter (150/30 μm outer/inner diameter, Part no. ES542, Thermo). We adopted a multinotch MS^3 method[48] to identify proteins based on MS^2 fragmentation and quantify them based on MS^3 measuring the reporter ions. The survey (MS^1) scans were performed at 120,000 FWMH resolution in the Orbitrap analyzer (Max IT, 120 ms; AGC target, 4×10^5 counts; microscans, 1; dynamic exclusion, 30 s). Using DDA, the precursor ions were isolated with a 1.2 Da window (±0.6 Da) and fragmented in the collision-induced dissociation (CID) cell in helium gas at 35% normalized collision energy (NCE). The tandem mass spectra were acquired in the ion trap (AGC target, 1×10^4 counts; max IT, set to auto mode; microscans, 1). The top 10 MS^2 fragments were isolated via synchronous precursor selection and dissociated further with higher-energy collisional dissociation (HCD) in nitrogen at 65% NCE. The generated MS^3 spectra were

acquired at 15,000 FWMH in the Orbitrap analyzer (AGC target, 5×10^4 counts; max IT, auto mode; microscans, 1).

Data Analysis. The primary MS−MS^n data were processed in Proteome Discoverer version 2.2, Thermo Scientific) with SEQUEST searching against the *Xenopus laevis* proteome version 9.2 containing 46,000 entries downloaded from Xenbase[12] and the mRNA-derived PHROG database[183]. The search parameters were as follows: enzyme, trypsin; the number of missed cleavages, maximum 2; fixed modification, cysteine carbamidomethylation, and TMT label; variable modification, methionine oxidation; minimum number of unique peptides, 1; mass tolerance for precursor masses, 10 ppm; and mass tolerance for MS^2 fragment masses, 0.6 Da. Peptide and protein matches were filtered to <1% false discovery rate (FDR), computed against a reversed-sequence decoy database. Common protein contaminants were manually removed. Unique and razor peptides were used for quantification. Proteins were quantified based on their MS^3 reporter ion abundances. The resolution of chromatographic peaks was measured in full width at half- maximum (fwhm) using MZmine version 3.0.21[186] with the Automated Data Analysis Pipeline for the chromatogram builder. GO annotation was performed in the PANTHER classification system version 16.0.[187]

Safety. Chemicals and biological samples were handled with care following standard safety protocols. Personal protective equipment was used to handle chemicals and sharp items. Forceps, needles, and nanoflow electrospray ionization emitters, which pose a needlestick hazard, were handled with care.

Scientific Rigor. We combined mathematical modeling and experimental validation to select protein ratios between the sample and YD carrier channels (see Results and Discussion). To

support YODEC development and validation, we eliminated biological variability (between embryos) by pooling tissue biopsies and embryos. In each part of the study, the proteomes were processed at the same time and by the same person to minimize variability in the sample preparation. YODEC was validated against a sample prepared using a nondepleted carrier (Control), the closest reference approach for this study. The FACS-isolated V11 cells were used to demonstrate proof-of- principle for different proteome amounts and conditions of sample collection. All experiments were conducted in 3 replicate measurements to aid statistical data interpretation. Precursor protein abundances were $\log_{10}$-transformed and median-normalized between the *Xenopus* and HeLa digests. The TMT reporter ion intensities in the analytical channels were $\log_{10}$-transformed and mean-normalized for comparison between the control and carrier experiments.

Data Repository. The MS data were deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[188] with the data set identifier PXD044787 and 10.6019/PXD044787.

2.4 Results and Discussion

The Rationale for "Dilute to Enrich". As our pilot experiments revealed that classical yolk platelet isolation is difficult to downscale, we proposed YODEC to be able to improve detection from limited proteomes by diluting the abundant yolk interferences. The yolk platelet pellet obtained from a limited cell population was too small for us to aspirate the YD supernatant with high recovery and reproducibility, without disturbing the pellet (**Figure 2.1**).

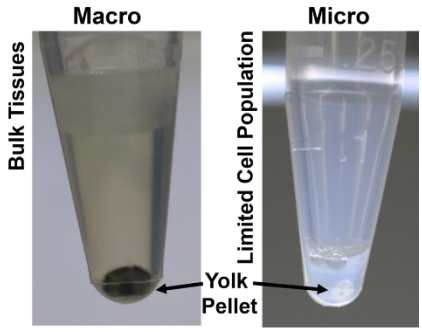

Figure 2.1. Yolk platelets were pelleted by centrifugation. Compared to tissue biopsies (macro image, left), limited cell populations (micro image, right) yielded yolk pellets that are difficult to identify for accurate isolation of the deyolked supernatant.

In our YODEC approach, illustrated in **Figure 2.2**, the sample proteome digest containing the abundant interferences is spiked with the same proteome depleted of the interferences, using isobaric tagging to encode the sample identity. Multiplexing barcoding also permitted relative quantification across multiple analyte channels (e.g., 4-plexing shown). Building on the recent improvements from the use of a carrier proteome,[36] we adopt the terminology of "carrier" TMT channel to mark interference-free/reduced proteome, albeit with use in a notably different context. Building on the previous applications, where the analyte and carrier proteomes were identical, our approach uses a carrier proteome that is intentionally different from that of the analyte; the carrier is depleted of abundant yolk platelets in YODEC. Because the depleted carrier proteome can be prepared by pooling and scaled for addition to the sample proteome, we

envision this “dilute to enrich” strategy to be generally scalable among different proteome amounts and specimens.

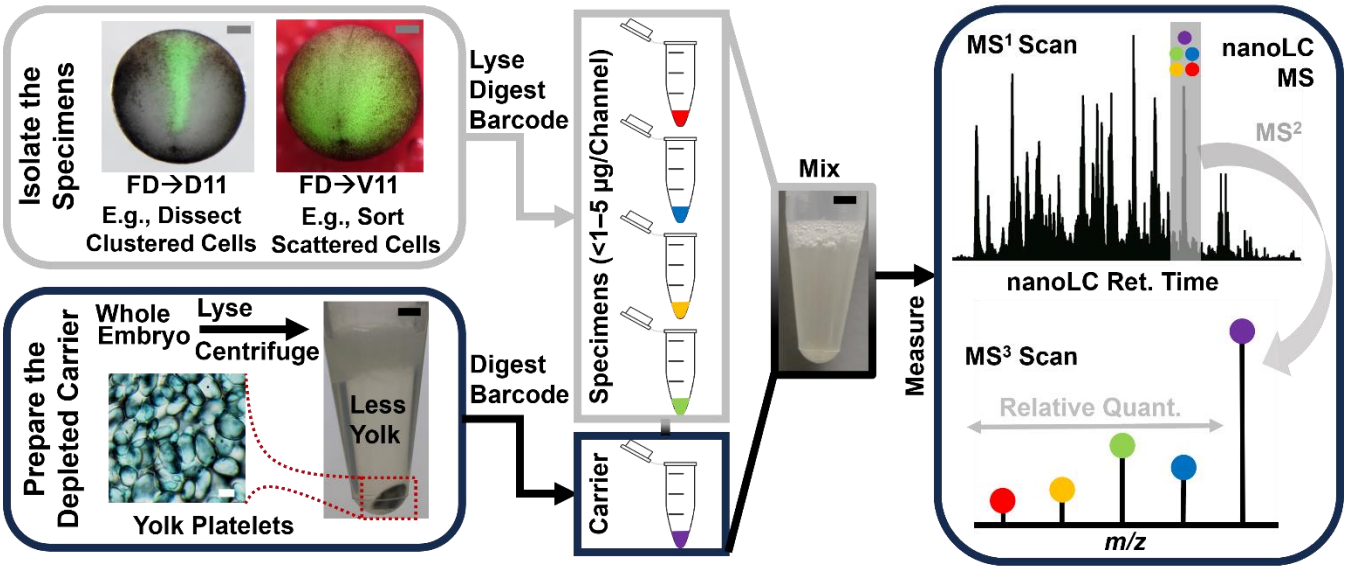

Figure 2.2. Our experimental strategy of “dilute to enrich” using a yolk-depleted (YD) carrier (YODEC). The example shows the fluorescent dextran (FD) labeling of identified neural-tissue-fated D11 and epidermal precursor V11 cell lineages in the X. laevis embryo (stage 13). The tight cluster of FD-labeled D11 cell lineage was dissected. The FD-labeled V11 clone, which scatters the embryo surface, was isolated through fluorescence- activated cell sorting (FACS). The collected proteomes were processed and TMT-tagged, then mixed with an isobarically tagged proteome digest prepared from a YD proteome. The samples were analyzed using nanoLC-HRMS employing multinotch precursor isolation and MS3-level quantification. Scale bars, 250 µm (gray); 2 mm (black).

The *X. laevis* proteome served as an ideal model to develop and validate the approach for limited amounts of proteomes, where the yolk is an omnipotent source of food for development and trouble for MS proteomics. **Figure 2.2** demonstrates YODEC on the dorsal-animal midline (called D11) cell, which normally develops into neural tissues and the epidermally fated V11 cell. These cells were identified based on pigmentation, morphology, and established cell fate maps in the 16-cell (stage 5) embryo (Experimental Section). These cells were microinjected with fluorescent dextran to label their descendent cell clone through gastrulation at stage 13 (Experimental Section). The offspring cells of the left and right D11 blastomere formed a tight

cluster, which could be isolated by manual dissection (as described elsewhere[14]). In contrast, cells of the left and right V11 lineages scattered the embryo too broadly and required fluorescence-activated cell sorting. A population of ~80,000 cells were collected from ~20 embryos in which both the left and right cells were labeled, and these cell populations were diluted to ~8,000 cells to test analytical performance for limited proteomes that would typically be available in a biological study. The YODEC channel was prepared from YD embryos (~100 μg of YD protein amount from ~5 pooled embryos), barcoded differentially, and mixed with the analytical channels (dissected D11 or sorted V11 cells, 1−5 μg total proteome; see design in **Table 2.1**, *Methods*). The resulting samples were measured by nanoLC-MS executing multinotched MS^3-level quantification (Experimental Section).

Severe Interferences within the *X. laevis* Proteome. To assess the performance of YODEC, we began by defining the analytical challenges that have so far hindered deep proteomics in early *X. laevis* development. We chose to benchmark proteomics in the yolk-laden *Xenopus laevis* against a commercial HeLa proteome, as this reference proteome is not known to suffer from as severe interferences. The *Xenopus* digest was prepared from neurula-stage embryos via standard bottom-up proteomics. The HeLa and *Xenopus* proteome digests were mixed at equal amounts based on a total peptide assay. The HeLa proteome (alone) and the HeLa-*Xenopus* mixture proteome were each analyzed to assess the effect of major yolk background on HeLa proteome detection.

The two detected proteomes were compared. **Figure 2.3A** maps the linear dynamic range of the detected proteomes based on label-free quantification (LFQ). The sample yielded 2,127 *X. laevis* proteins compared to 5,313 from HeLa. The amphibian proteome spanned a 100 times wider

quantitative dynamic range based on the calculated LFQ indices. Vitellogenin (Vtg) proteins, the food source of the early developing embryo,[13] encompassed 80% of the total proteome based on LFQ, with Vtga2.L and Vtgb1.L ruling the majority (~50% of the total proteome).

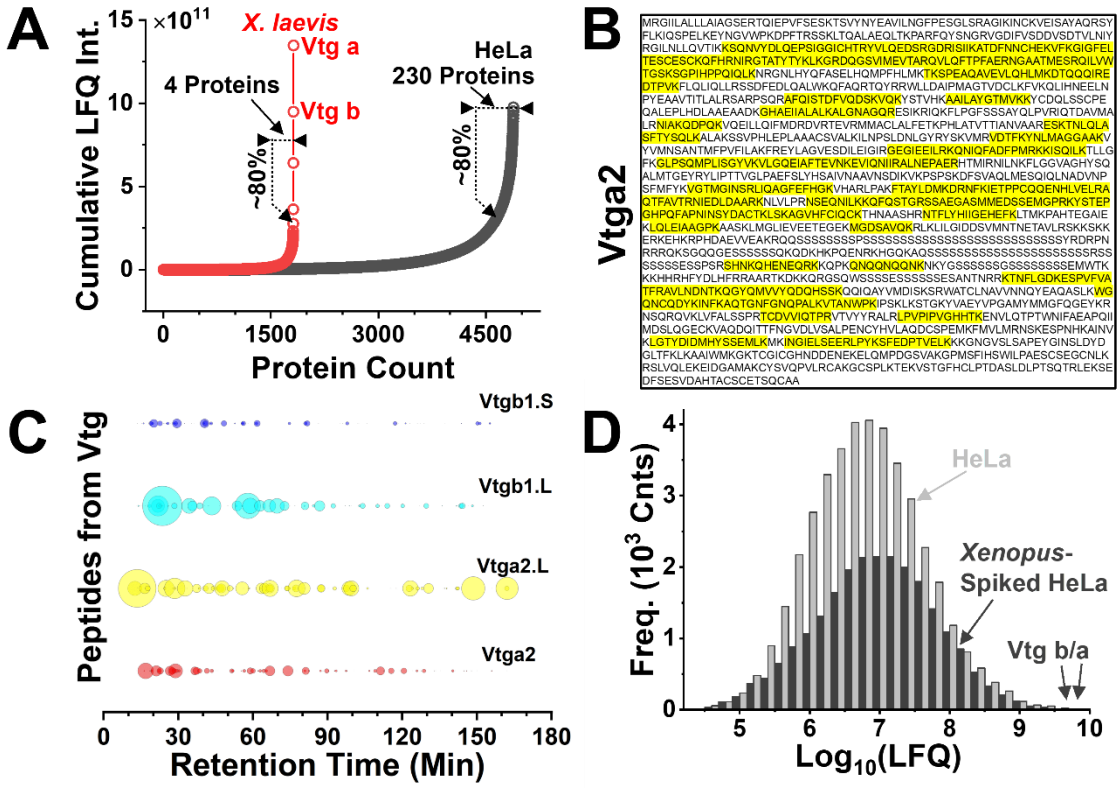

Figure 2.3. Interferences from abundant proteins. **(A)** Cumulative protein abundances in X. laevis embryos and the reference standard HeLa, revealing vast quantitative differences in the proteome distribution. Compared to HeLa, the dynamic concentration range of the X. laevis proteome was 100 times wider based on LFQ and was dominated by vitellogenin (Vtg) yolk proteins. **(B)** Example of Vtga2 yolk protein detection based on the sequencing of 101 experimentally identified proteotypic peptides (yellow). **(C)** Example of yolk peptides eluting over the entire chromatographic separation. The size of each circle indicates the relative signal abundance among the peptide ions. **(D)** Comparison of the dynamic range of quantification between HeLa and Xenopus-spiked HeLa proteomes following median normalization.

Even long after yolk consumption had commenced to feed embryonic development, the larval proteome was still dominated at ~70% by these few yolk proteins based on our experiments in this study. For comparison, the same 80% of the most abundant portion of the proteome was contained by ~230 proteins in HeLa (**Figure 2.3A**). As this two-proteome model recapitulated

vast differences in protein concentration distribution between the species, it offered a reliable reference to develop and validate YODEC.

These copious Vtg-derived peptides were anticipated to challenge MS proteomics in many ways. Tryptic digestion *in silico* (PeptideMass, Expasy) predicted that each isoform of Vtg would produce ~100 peptides on average. Experimentally, our data set identified ~450 peptides from four isoforms of Vtg. For example, 101 detected peptides were mapped to Vtga2, thus explaining the high 60% sequence coverage for this protein (**Figure 2.3B**). As traced in **Figure 2.3C**, the yolk peptides formed abundant signals over practically the entire range of the chromatographic gradient. We tested whether this large peptide background affected the separation performance. The distribution of chromatographic peak widths was ~3-fold wider in the control than in the YD proteome digest on the same LC- HRMS instrument. Although the *Xenopus*-spiked HeLa yielded a comparable number of HeLa proteins to the HeLa-only control (5,313 vs 5,474 proteins, respectively), there were quantifiable differences in analytical performance. Up to ~15% of all peptide spectral matches (PSMs) in *Xenopus*-containing HeLa were from Vtg. Spiking of the HeLa proteome with *Xenopus* led to 30% fewer PSMs. As shown in **Figure 2.3D**, the detectable proteomes were normally distributed, with less proteome diversity quantified across the comparable linear dynamic range of quantification in *Xenopus*. These results illustrate that the abundant yolk proteome impedes identification and quantification.

Classical yolk depletion is prohibitively difficult to scale to limited cell populations. From whole embryos, centrifugation of the platelets[14,45] allowed 2,836 proteins to be profiled vs 1,453 proteins without. These results aligned with our previous doubling of proteins that were detectable in the developing neural ectoderm tissue at stage 10 using yolk depletion. Yolk

depletion, however, is not exhaustive.[14] From whole embryos, we found the approach to still preserve yolk proteins amounting to ~25% of the total proteome by LFQ. This result fairs well with our previous finding of ~30% yolk remaining after depletion of the yolk platelets.[14] As discussed earlier, these abundant yolk contaminants hamper proteome coverage. Another difficulty with the approach of physical depletion is due to the manual finesse that is necessary to scale to limited proteomes, such as developing single cells and cell lineages in *X. laevis* (partially reviewed in the literature[53]). For a tissue equivalent to ~5% of the embryo, which estimates to the D11 and V11 lineage we target in this study, the yolk platelet was barely visible after sucrose centrifugation (**Figure 2.1**), a step necessary for yolk depletion in high fidelity.

YODEC Is Convenient and Scalable. With only a single depletion step needed to prepare the depleted carrier proteome, YODEC simplified experimentation and improved scalability toward lower amounts of starting proteomes; once prepared, the YD carrier proteome could be spiked into any amount of the limited proteome. We set out to define the practical boundaries of the method based on mathematical modeling of dilution as the analyte proteome is mixed with the YD carrier proteome. Taking the LFQ abundances from our previous study to initiate the model,[14] **Figure 2.4A** simulates the fractional composition of the proteome that would result upon spiking the embryonic proteome with 1 to 20 times equivalent YD carrier. In the case of multiplexing quantification, this ratio was calculated by dividing the amount of proteome contained in the carrier channel by the total proteome resulting from all the analyte channels. The experimental design is presented in **Table 2.1** (*Methods*). As predicted, the addition of the YD carrier helped dilute yolk interferences (Vtga2 and Vtgb1), while enhancing signal abundance for analytes not associated with the platelets (e.g., Stt3b, Rpl9, Nfs1). The calculation forecasts a plateauing in these respective improvements starting with carrier loads of ~4 times the analytical

channels. We considered this ratio as the boundary condition for experimental testing in this study.

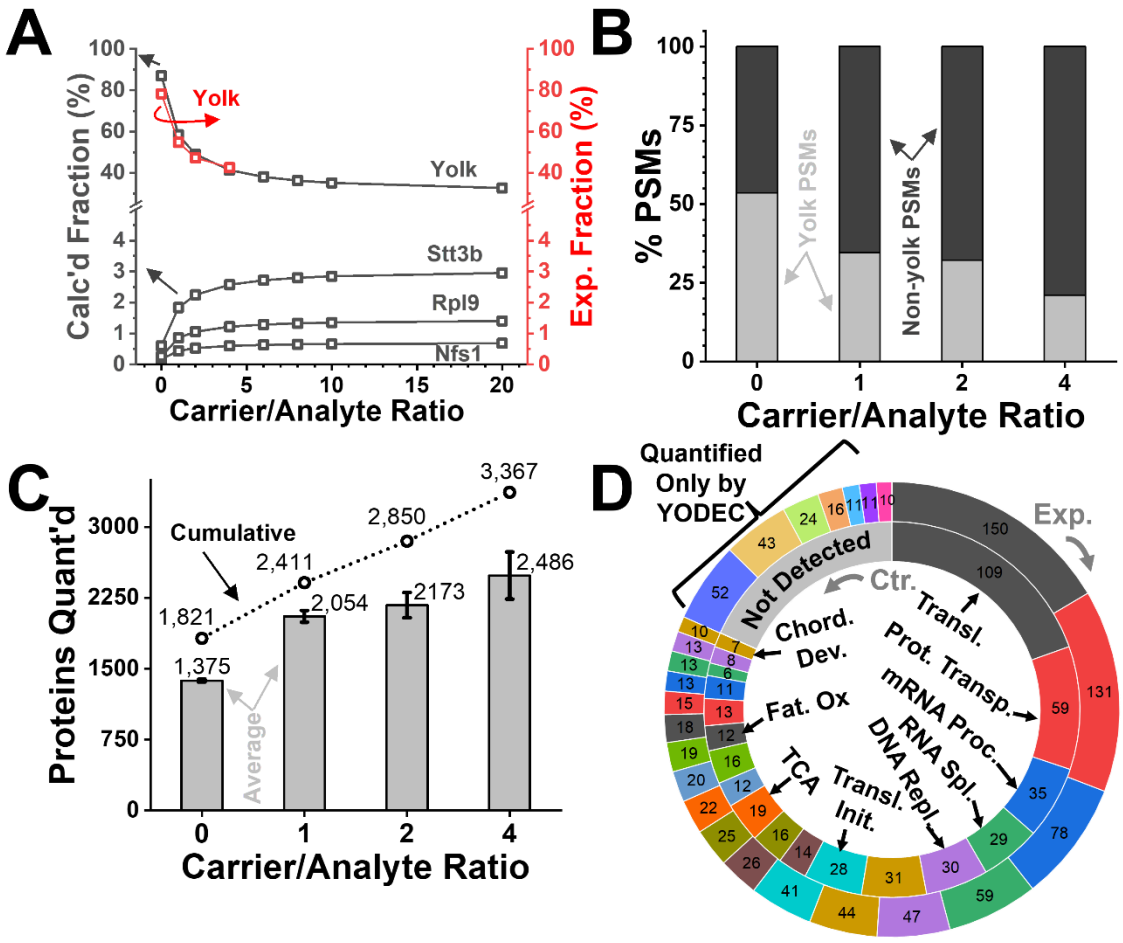

Figure 2.4. YODEC improves proteome coverage and quantification. (A) Fractional composition of the X. laevis proteome upon progressive dilution with the yolk-depleted proteome from in silico modeling (black) vs experimental testing (red). For a proteome mixture containing TMT- barcoded X. laevis proteome at a 1.5:1:1:1.5 analytical ratio, increasing carrier proteomes (B) reduced PSMs on yolk peptides and (C) improved the number of quantifiable proteins. (D) Comparison of GO assessment of molecular function for proteins that were quantified without (Control, inner circle) and with YODEC using a 4-fold carrier load (Experimental, outer circle). Close-up in Figure S2. Identical GO categories are coded in the same color between the control (inner circle) and experimental (outer circle) cases.

We studied the model with multiplexing barcoding quantification. For this portion of the study, the D11 lineage was selected, because its descendent cells form tight clusters in the embryo that

can be accurately dissected following our established protocols.[54] **Figure 2.2** presents the fluorescent labeling of the clone in the neurula (stage 13) embryo. The YODEC carrier was prepared from a pool of whole embryos at the same developmental stage. To facilitate method testing, we removed biological variability in this portion of the study by analyzing in technical replicates (n = 3) a shared reference proteome digest (yielding ∼1−5 μg per experiment). To establish quantitative performance, the control sample proteomes were isobarically barcoded and mixed at a 1.5:1:1:1.5 ratio (using the TMT channels 126, 127, 128, 129), then spiked with a differentially TMT-tagged YD carrier proteome digest (TMT channel 130) at an equivalent, double, or quadruple amount of the net proteome among the analyte channels. Results from the experimental data (**Figure 2.4A**) confirmed the dilution of the total yolk to a ∼40% fraction of the total proteome while enhancing the signal for other peptides, such as Stt3b, Rpl9, and Nfs1, which we randomly selected as examples. As traced in **Figure 2.4A**, the experimental protein concentrations agreed with our theoretical predictions. Dilution with the YD carrier improved the economy of PSMs for sequencing nonyolk peptides. Based on these results, we ascribed the basic working principle of YODEC to concentration dilution.

Protein identification and quantification improved with YODEC. Between the technical replicates, the cumulative identifications improved from 2,053 (Control, without YODEC) to 4,285 proteins using YODEC with a 4-fold carrier load. **Figure 2.4C** shows the number of proteins that had quantifiable values in all analyte channels with increasing amounts of the YD carrier.

As shown in **Figure 2.4D**, the proteins carried out a broad range of molecular processes and biological functions (see *Appendix* for close-up images). They participate in pathways controlling the cell cycle, cell division, RNA processing and splicing, mitochondrion

organization, respiratory chain complex I, microtubule-based processes, actin polymerization, and nucleocytoplasmic transport. Some of these proteins were only detectable using YODEC (compare inner vs outer circles). Ontology (GO) annotation found no detectable differences in the subcellular distribution of proteins that were quantified in the control vs YODEC (4× carrier) experiment (see *Appendix* for close-up image).

Figure 2.5 visualizes the overlaps between the control and experimental conditions (1, 2, and 4× carrier). YODEC using a 4-fold carrier load provided the best overall performance. The percentage of proteins that were quantified without missing values in any of the analytical channels remained consistent at ~80% between all the experiments.

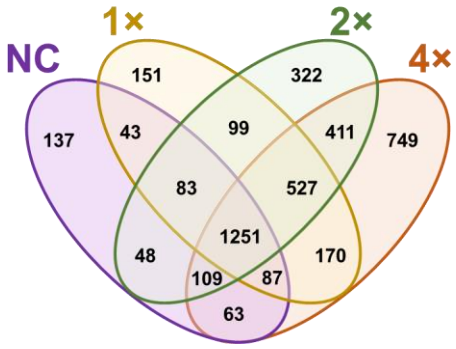

Figure 2.5. Overlap of proteins quantified by YODEC at different carrier amounts.Analyte channels were mixed with the carrier tagged proteome at 1×, 2×, and 4× the different carrier/analyte ratios. No carrier (NC) served as the reference. A 4-fold carrier helped quantify 749 proteins uniquely.

These data suggested improved quantification for low- abundance proteins. YODEC allowed us to quantify several proteins that had significantly lower abundance than the distribution mean based on the calculated, $\log_{10}$-normalized, and mean-centered data (**Figure 2.6**). Furthermore, increasing YODEC carrier loads maintained good analytical reproducibility, with the accuracy of relative quantification averaging to ~12.5% error for 1:1 and 1:1.5 proteome mixtures (**Figure**

2.7). These results demonstrated deepened quantitative proteome coverage with good quantitative robustness by YODEC.

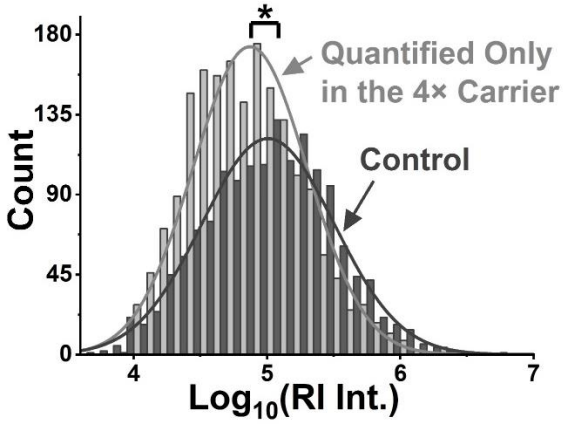

Figure 2.6. Protein quantification by the control vs. YODEC. The reporter ion signal abundances were log10-normalized. The control and the YODEC data were median-normalized. Using a 4-times deyolked carrier proteome, YODEC was able to quantify proteins that were undetected in the control. These newly quantifiable proteins populated the lower concentration range of the proteome.

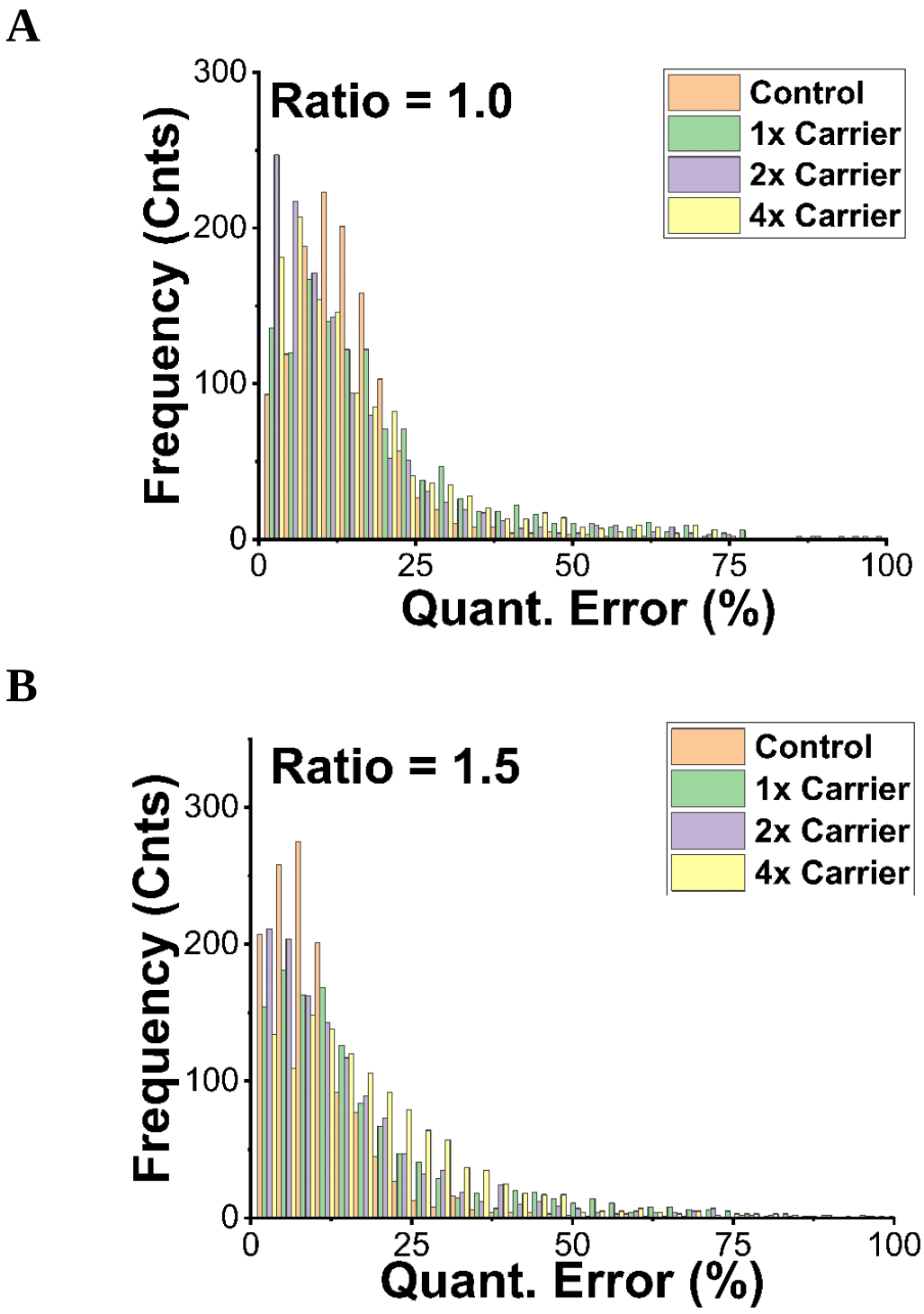

Figure 2.7. Accuracy of quantification in YODEC.The % error of quantification was computed as the ratio between observed (measured, experimental) and expected (mixed, control) proteome mixtures. The TMT ratios were tested for analyte channels containing (A) an equal amount and (B) 1.5-fold different. The DY carrier proteome was spiked into these samples at varying (1×, 2×, 4×) ratios to the total proteome over the analyte channels.

YODEC Dilutes to Enrich. The mechanism underlying "dilute to enrich" employs notable differences from the carrier proteome approach of SCoPE MS. Taking YODEC as an example, we rationalized that mixing the *X. laevis* proteome with its YD proteome would systematically improve the detection/quantification of nonyolk proteins, while diluting proteins contained in yolk platelets, including yolk. In contrast, dilution of the HeLa proteome digest with itself (without depletion of any proteins) would not alter the respective makeup of the proteome as a result of the carrier approach. Multivariate data analysis of MS^1 abundances of proteins (*Xenopus* or HeLa) between the control and experimental conditions helped to objectively assess these scenarios. **Figure 2.8** presents the hierarchical cluster and heat map analysis of the ~50 proteins that were statistically most significantly changed in abundance with dilution of the embryo proteome by the YD carrier in YODEC (**Figure 2.8A**) or the HeLa proteome in the carrier proteome approach (**Figure 2.8B**). Close-ups of these figures are available in the *Appendix*. The technical triplicates formed data with tight clustering for each different dilution, assuring us of sufficient reproducibility to facilitate the results interpretation.

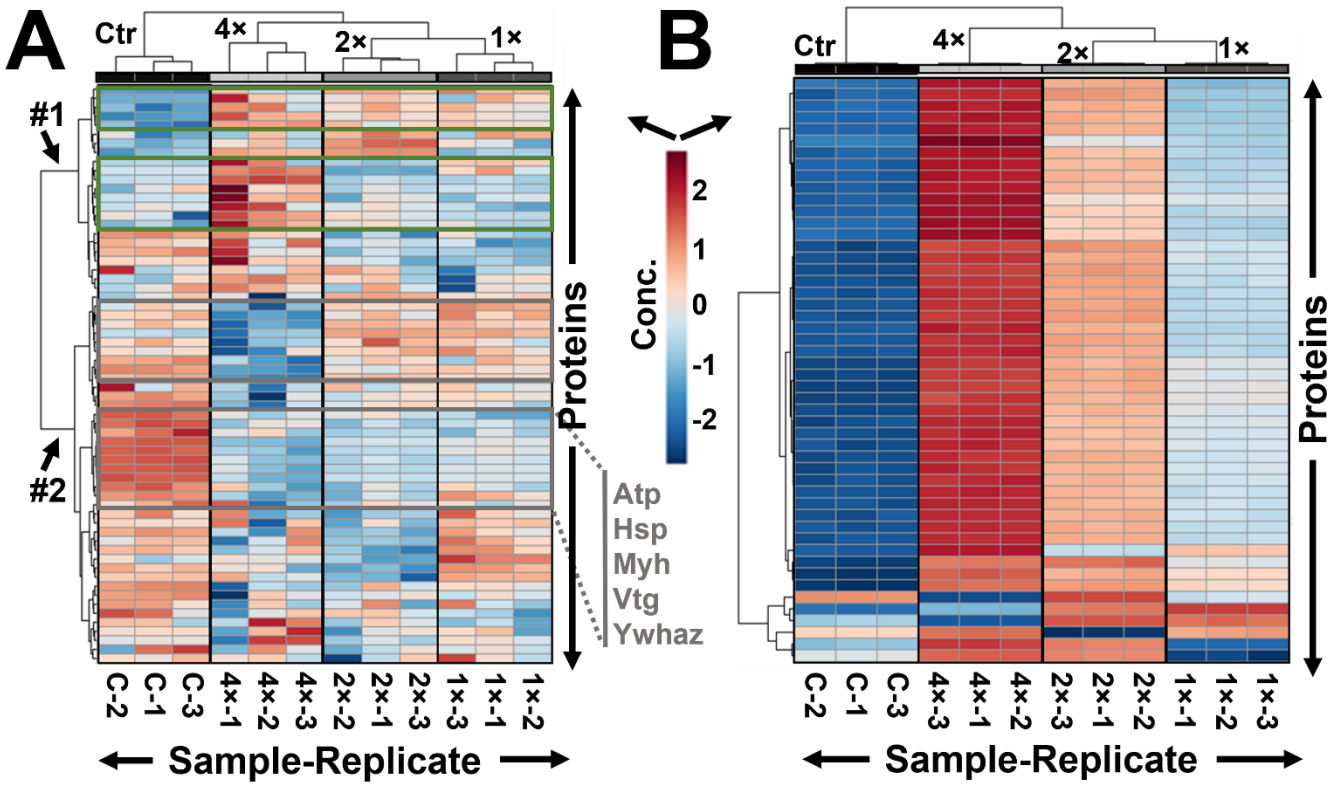

Figure 2.8. Mechanistic comparison of the "dilute to enrich" approach using the X. laevis and HeLa proteomes. Heatmap and unsupervised cluster analyses on the 50 statistically most significantly differential proteins between the control (C) and experimental conditions using carrier proteomes in 1-, 2-, and 4-times excess (samples labeled 1×, 2×, and 4×, respectively). Each sample was analyzed in three technical replicates (sample labels appended with −1, −2, or −3). These measurements revealed **(A)** differential proteome enrichment during YODEC, where the X. laevis proteome was mixed with its yolk-depleted proteome (yielding Groups #1 and #2), **(B)** but not in the carrier proteome (SCoPE) strategy on HeLa, where protein amounts were uniformly increased. Close-ups of these plots are available in see Appendix. Key to color code: blue, signal loss; red, signal enrichment. Sample labels specify the experiment (control vs YODEC carrier) and the number of technical replicates.

The proteomes responded strikingly differently to isobaric addition. YODEC led to a differential proteome enrichment in *X. laevis* (**Fig. 2.8A**). The responses of the proteomes were strikingly different to the isobaric addition. YODEC led to a differential proteome enrichment in *X. laevis* (**Figure 2.8A**), whereas the proteome carrier approach indiscriminately enriched all proteins in HeLa using the carrier proteome approach (**Figure 2.8B**). Close-ups of these plots are available with gene labels in *Appendix*. Cluster analysis of the data found the *Xenopus* proteome to have

responded to spiking in two main ways, forming protein Groups labeled #1 and #2, with smaller subgroups within each (labeled in **Figure 2.8A**, close-up in *Appendix*,). Proteins in Group #1 decreased in concentration upon isobaric addition. Representative members were vitellogenins (Vtga1, Vtga2, Vtgb1, and Vtgb2), the main content of yolk platelets,[14] and other proteins known to associate with yolk platelets in this species,[13] including Atp5a1, Eef1g, Hadhb, Hsp90b1, HspA8, Hspd1, Myh9, Pdia3, Rheb, Rhoa, Vdac2, and Ywhaz. In contrast, proteins in Group #2 underwent progressive enrichment with increasing carrier loads. These proteins, most occupying the lower domain of the concentration range, are known transcription factors or components of signaling pathways of neural tissue genesis, including Actr2, Csnk2b, Ctnnb1, Eno1, Itgb1, Mapk1, Ocln, Pgm1, Psmd14, Sbds, Serpinb6, Sf1, and Smarca5. In stark contrast, experiments on the HeLa proteome revealed no improvement in proteome coverage, except for monotonously increasing concentrations for nearly all proteins with the addition of the carrier channel (**Figure 2.8B**, close-up in *Appendix*). Combined, these results agreed with protein dilution resulting from platelet isolation in YODEC and demonstrated mechanistic differences from the SCoPE carrier.

We further benchmarked YODEC against the standard carrier approach for *Xenopus* proteomics. We compared the number of proteins that were reproducibly quantifiable across analyte channels in a multiplexed quantitative study. We differentially barcoded and mixed the D11 proteome at a ratio of 1.5:1:1:1.5 following our earlier design (see *Methods,* **Table 2.1**) that validated quantitative accuracy (**Figure 2.7**). The standard shotgun workflow (Control) was evaluated against YODEC applying 4 times the YD carrier load than the analytical channels combined. As a reference for the conventional proteome carrier approach from SCoPE,[36] the TMT-multi-plexed sample was measured in the presence of 4 times the amount of its barcoded self as the carrier proteome (without yolk depletion). The results of this study are summarized in **Table 2.2.**

Table 2.2. Benchmarking of YODEC for protein quantification without missing values on a limited amount of the *X. laevis* proteome

Carrier	Analytical Method	Yolk Content (%)	Proteins Identified	Proteins Quantified
None	Control	73	4,772	4,305
Embryo as Carrier	SCoPE	76	3,827	3,334
Yolk-Depleted Carrier	YODEC	40	5,867	5,218

YODEC was able to quantify 5,218 proteins consistently among the replicates in a notable improvement from the reference methods. A comparably high (>70%) measured yolk content gave rise to ~23% and ~28% of the PSMs to be acquired on yolk peptides in the control and SCoPE carrier approach, respectively. In comparison, a reduced, ~40% yolk content in YODEC halved the number of PSMS on yolk peptides to ~10%, thus freeing up valuable bandwidth to sequence (MS^2 stage) and quantify (MS^3 stage) non-yolk proteins. The reference carrier approach (SCoPE) maintained high yolk content and disadvantaged protein quantification like the Control, whereas YODEC leveraged the YD carrier to effectively dilute yolk for improved protein quantification.

YODEC Scaled Well to Limited Proteomes. To test the utility of the approach, we applied YODEC on a limited population of embryonic cells. For this part of the investigation, we chose the epidermally fated lineage that the V11 cell forms from the 16-cell embryo. Unlike the previously analyzed tight clusters of the neural-tissue-fated D11 lineage, which could be readily dissected as a tissue, these cells scatter the surface of the embryo too broadly for biopsy. As earlier with the D11 clone, we fluorescently labeled via microinjection the V11 clone from the cleave-stage embryo to the neurula (stage 13, **Figure 2.2**) before gently dissociating the embryo into a cell suspension containing BSA to reduce cell adhesion and employed FACS to isolate ~8,000 cells, containing ~1.5 μg of total proteome (including yolk and BSA contaminant).

Whole embryos at the same developmental stage were yolk-depleted to prepare the YODEC carrier.

The depth and robustness of proteome quantification were established. To test quantitative accuracy, the V11 proteome was divided and barcoded to prepare a mixture containing the proteome digest at a 1:1.5:1 (known) ratio. This analyte was measured using the standard shotgun proteomics or mixed with 4 times the amount of a YD carrier (see design in **Table 2.1**, *Methods*). Optionally, the YODEC samples were fractionated at high pH before low-pH analytical HPLC in an effort to improve the coverage. Protein quantification is compared between these methods in **Figure 2.9A**.

While the standard method identified 252 proteins (207 quantified), YODEC was able to report on 1,427 proteins (1,117 quantified). Eight-step high- pH fractionation expanded identifications to 3,818 proteins (3,133 quantified), of which 2,175 proteins were only quantifiable using this strategy. Notably, these YODEC- afforded 5−15-fold improvements in the depth of the quantifiable proteome were achieved simply on the principle of mixing analytes rather than laborious or specialized sample preparation or instruments.

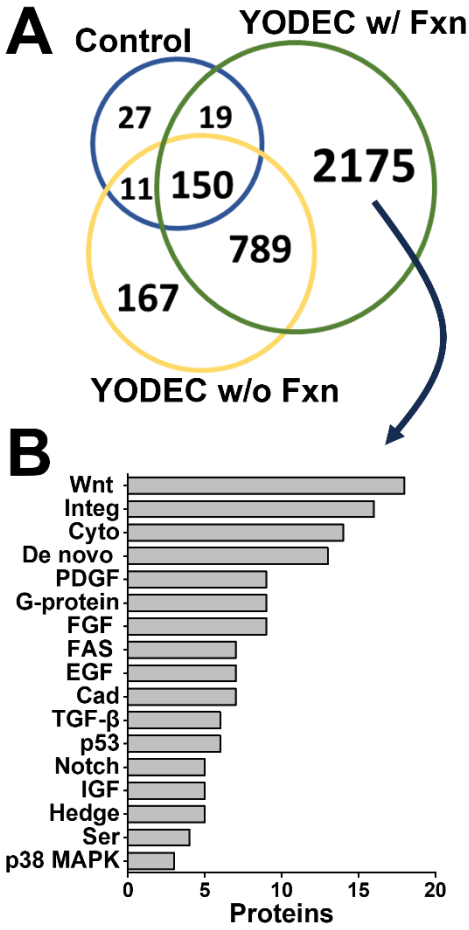

Figure 2.9. Application of YODEC to limited amounts of proteomes available from ~8,000 sorted cells (~1.5 µg total, or ~150 ng non-yolk proteome). **(A)** Improvement of protein quantification from the standard approach (Control) through the integration of YODEC with fractionation (Fxn). **(B)** Representative signaling pathways enriched by proteins that were quantifiable only by YODEC with fractionation. Key: Cad, cadherin signaling; Cyto, cytoskeletal regulation by Rho GTPase; De novo, de novo purine and pyrimidine synthesis; EGF, epidermal growth factor signaling; FAS, Fas signaling; FGF, fibroblast growth factor signaling; G-proteins, Heterotrimeric G-protein signal- ing pathway-Gi alpha and Gs alpha mediated pathway; Hedge, hedgehog signaling; IGF, insulin like growth factor signaling; Integ, integrin signaling; Notch, notch signaling; p38 MAPK, p38 MAP kinase pathway; p53, p53 signaling pathway; PDGF, platelet derived growth factor signaling; Ser, serotonin mediated signaling; TGF-β, transforming growth factor-β signaling; Wnt, Wnt signaling.

We explored the canonical functional roles of the proteins that were detectable only using YODEC. As shown in **Figure 2.9B**, the 2,175 proteins that were only quantifiable by YODEC with fractionation were enriched in signaling pathways related to embryonic cell differentiation and tissue induction. For example, Wnt signaling plays a critical role in guiding cell fates and

axis formation within the embryo. Integrin, hedgehog, notch, and EGF pathways are important for the proliferation and cell differentiation of embryonic cells, whereas cytoskeletal regulation by Rho GTPase and cadherin signaling is known to control cell migration and movement during embryonic development. Without YODEC, these signaling pathways would have been undetected using the classical approach of bottom-up proteomics, sucrose-centrifugation-based yolk depletion, or the proteome carrier approach (SCoPE). The ability to detect these proteins in such pathways opens the potential to better understand cell differentiation, e.g., induction of the ectoderm into neural tissue.

2.5 Conclusions

In this study, we introduced the idea of a "dilute to enrich" approach to improve proteome coverage from a limited amount of proteome. Although continual analytical advances improved sample collection, preparation, analysis, and detection strategies, abundant proteins still present a fundamental limitation to sensitivity due to interference at nearly every step of the analytical workflow. We designed YODEC to address this challenge in the yolk-laden *X. laevis* embryo by spiking the endogenous proteome with a YD proteome carrier. This carrier was prepared by the macroscopic depletion of yolk platelets from the cells. Although YODEC uses isobaric signal boosting like the SCoPE carrier proteome[36] to improve peptide detection sensitivity, we found that the mechanism of signal improvement by YODEC is fundamentally different. YODEC employs differential concentration dilution to remove abundant protein interferences, while concurrently also enhancing sensitivity via the mechanism of isobaric signal enhancement as pioneered in the SCoPE carrier approach. Using the HeLa reference proteome, we validated that YODEC works by diluting interferences from yolk signal faster than less abundant proteins while enhancing or preserving quantitative performance.

Technical simplicity secured experimental scalability for deeper proteomics of limited populations of cells. In the *X. laevis* neurula, YODEC reduced a massive ∼70−90% yolk proteome content to 20−40% based on LFQ in a dissected neural-tissue-fated D11 lineage. In combination with high-pH fractionation, the approach enhanced the number of identifiable proteins by 15-fold, permitting the quantification of 3,133 proteins in the epidermally fated V11 cells from FACS isolation. A collection of 2,175 proteins that were only detected by this approach were enriched in signaling pathways of critical importance to cell differentiation, tissue induction, and patterning of the embryonic body plan.

While we demonstrated the utility of YODEC in stage-13 *X. laevis* tissues, the approach can be tailored and extended to other proteomes whence abundant proteins can be isolated through physical (e.g., manual methods) or chemical (e.g., antibody-based depletion) means to "dilute to enrich". Of course, different proteome backgrounds ideally call for optimization of the depletion approach and depleted carrier load to aid quantitative sensitivity. Further, in developing organisms, like the (*X. laevis*) embryos studied here, active proteome reorganization presents an opportunity to specifically tailor the depleted carrier proteome to the specimen with spatiotemporal resolution.[53] Pooling of tissues over multiple developmental stages to prepare the depleted carrier provides a simple solution to help adjust the YODEC carrier proteome. It also holds the potential to advance the study of post- translational modifications from limited samples by permitting enrichment for targeted modifications in the carrier proteome. YODEC explores a conceptual strategy that we anticipate being broadly adoptable to deepen the measurable proteome across the realms of the organism, organ, tissue, and cell.

Acknowledgments. Parts of this work were funded by the National Institute of General Medical Sciences (award 1R35GM124755 to P.N.), the University of Maryland-National Cancer Institute

Partnership Program (seed award to P.N.), and the COSMOS Club Foundation (fellowship award to L.R.P.) We thank Ken Class for assistance with FACS at the MPRI Flow Cytometry and Cell Sorting Facility. The content of the work reported here is the sole responsibility of the authors and does not necessarily represent the official views of the funding sources.

Chapter 3: Microscale Sample Processing for Limited *Xenopus* Embryonic Tissues

This chapter is based on the Pade, L.R. and Nemes, P. (2024) manuscript draft in preparation.

Contribution: L.R. P. collected and prepared the samples for proteomic analysis, analyzed the samples via LC-MS, performed data analysis, interpreted the results, drafted the manuscript, and revised the contents.

3.1 Abstract

Advancements in MS over the years have been crucial to the field of proteomics, improving our knowledge of mechanisms underlying various biological processes influencing health and diseases. Nevertheless, achieving coverage comparable to genomic and transcriptomic studies remains a challenge for proteomics, primarily due to the unfeasibility of proteome amplification like their genomic counterparts. Sample losses due to multistep sample processing further challenges detection sensitivity especially from paucicellular samples. In *Xenopus* embryonic tissues their limited size and presence of abundant yolk proteins severely impact global proteome analysis. To improve the detection sensitivity, we developed a microscale yolk depletion device capable of isolating yolk platelets from low volumes of tissue lysates (≤ 10 µL), while minimizing protein losses during processing and precluding the need for pooling large tissue amounts. In parallel, we developed an LC-MS approach to measure the proteome in a discovery setting. We maximized protein identifications by employing a new-generation micro-pillar array column (µPAC), tailored in performance for the separation rate, solvent gradient, and column temperature. Next, we combined our microscale yolk depletion device with high-pH fractionation, followed by low-pH µPAC LC to process and analyze the embryonic tissues. This integrated approach enabled ~8,000 protein identifications from single dissections of *Xenopus*

tissues, providing 4-fold deeper coverage compared to our previous studies on bulk or pooled tissues.

3.2 Introduction

Gene products like proteins and metabolites are produced downstream of genes and transcripts and are known as effector molecules as they catalyze all reactions and drive signaling. With the advent of MS, analyzing these molecules with high sensitivity across all spatial and temporal domains is now feasible.[189] Several studies have highlighted the importance of directly measuring gene products as their expression does not always correlate with their genes or transcripts counterparts.[190] Especially, in dynamic systems such as the developing embryos uncovering the active state of the cells is critical to improving our knowledge of how morphogen gradients drive tissue induction and differentiation. A recent quantitative proteomic study in developing embryos comparing the proteomic profiles of two species *Ciona* and *Xenopus* dispelled an old model of conservation in species at early embryonic stages based solely on the genomic data. In contrast, their analysis revealed maximal proteomic differences early on especially during gastrulation and neurulation.[191]

The field of proteomics faces considerable challenges due to the lack of amplification technology available for genomics. Challenges in sensitively detecting proteins from ever decreasing cell populations in turn spurred the development of innovative strategies to advance MS based analyses. A few of the challenges include sample loss during processing, protein digest complexity, and a large concentration range of different proteins in the samples.[71, 192] Innovations pertaining to all aspects of the proteomic workflow encompassing sample collection and processing, digestion, separation, ionization, and detection proved useful in tailoring methods to specific biological questions at hand. While approaches such as multiplexing[193],

wide window data dependent acquisition[194], DIA acquisition[195], single cell proteomics (SCoPE)[180], and yolk-depleted carrier[196] (YODEC) use innovative ways to maximize MS duty cycle, improve detection sensitivity, several other strategies focus on reducing protein losses during processing and reducing sample complexity.

Strategies to confine processing volumes and limit sample transfer such as micro and nanodroplet processing platforms (microPOTS, nanoPOTS)[178, 197], filter-aided sample preparation (microFASP)[179], oil-air-droplet chips[198] permitted the processing of a small population of cells or single cells.[199] Reducing harsh detergents further eliminated the need for sample clean-up reducing the number of sample processing steps. For example, surfactant-assisted one-pot sample preparation coupled with MS (SOP-MS) leverages another MS-compatible surfactant, DDM, to further scale down this process to differentiate proteomic profiles of single cells from cancer tumors and primary metastases in a mouse model.[200]

In this study, we developed a sample processing strategy for limited amounts and volume-limited cell populations to enable the study of spatial heterogeneity in these developing cell populations as embryogenesis progresses. Although *Xenopus* is extensively used to understand processes involved in early embryonic development, regeneration, and diseases, proteomic studies in this model are largely limited due to abundant yolk proteins. Macroscale yolk depletion methods have limited the studies due to the need for pooling large tissue amounts.[183] To downscale the processing volumes and facilitate deep proteome coverage in a small population of cells we developed a microscale yolk depletion approach (μYODE) to preclude sample pooling and reduce sample loss. Combining μYODE with micropillar array separation to reduce sample complexity we identified over 8000 proteins from single dissections of neural cell clones (~1000

cells). We achieved one of the deepest proteomic coverages from single *Xenopus* tissue dissections (~200 ng yolk free protein amount) reported in the literature.

3.3 Experimental section

Materials and Reagents. LC-MS grade solvents and reagent-grade chemicals were purchased from Fisher Scientific (Pittsburgh, PA) or Millipore Sigma (St. Louis, MO) unless specified otherwise. Trypsin protease (MS grade) and green fluorescent dextran (FD, Alexa Flour 488, 3000 MW) were purchased from Fisher Scientific. Combretastatin A-4, cytochalasin D, nonidet P-40 substitute (NP-40), dithiothreitol (DTT), and iodoacetamide (IAD) were from Sigma-Aldrich (St. Louis, MO). Borosilicate capillaries were purchased from Sutter Instrument (Novato, CA).

Solutions. 100% (w/v), 50%, and 20% Steinberg's solutions (SS) were prepared according to established protocols for embryo culture and tissue dissection.[201] The conventional lysis buffer contained 1% (v/v) sodium dodecyl sulfate, 150 mM NaCl, 5 mM EDTA, and 20 mM Tris hydrochloride. The yolk depletion buffer consisted of 250 mM sucrose, 1% nonidet P-40 (NP-40), 5 mM EDTA, 20 mM Tris-HCl, 10 μM combretastatin 4A, and 10 μM cytochalasin D, and Tris-HCl, it was prepared following previously outlined protocols.[159, 183]

Animal Care and Embryology. A breeding colony of adult male and female *X. laevis* frogs purchased from Nasco (Fort Atkinson, WI) or Xenopus1 (Dexter, MI) was maintained and handled in accordance with humane protocols approved by the Institutional Animal Care and Use Committee of the University of Maryland, College Park (R-FEB-21-07). Natural mating was set up to obtain embryos using established protocols.[201] Embryos were treated with 2% (w/v) cysteine solution (pH 8) to remove the jelly coat and transferred to 100% SS solution. Embryos

with stereotypical pigmentation were sorted at the 2-celled stage and cultured to the 16- or 32-cell stage (Nieuwkoop and Faber, NF stage 5 or 6) in 100% SS.[8-9, 202] For method development, neural precursor cells (dorsal midline, D11) were identified based on stereotypical pigmentation and location in the embryo and microinjected with the 1 nL of 0.5% v/v fluorescent (Alexa Fluor 488) dye to trace their descendent lineages. Embryos were cultured to the beginning of gastrulation (NF Stage 10) in 50% v/v SS solution.

Tissue Collection. For developing the µYODE approach we chose to use neural-fated cell clones for ease of microinjection and collection of a large number of samples to test different conditions, additionally, the neural-fated cell clones were representative of the tissues under future spatial study. We manually dissected the neural-tissue-fated cell clone at Stage 10 using sharp forceps in 50% SS in multiple replicates. Single dissections obtained from one embryo were individually frozen in tubes at -80 °C until further processing. Similarly, for the spatial study, single dissections of the SMO and the NE tissues were frozen in individual tubes.

Sample Preparation. Capillary tubes were prepared by cutting the tubes to the desired length and sealing one end of the tube using a Bunsen burner. For non-yolk depleted conditions, dissected tissues were processed either in a LoBind Eppendorf tube or sealed capillary using standard bottom-up proteomics protocol. Alternatively, for yolk-depleted samples, yolk depletion was performed in an Eppendorf tube or glass capillary to separate yolk platelets via centrifugation using the yolk depletion buffer.[159] The yolk-depleted supernatant was then processed using standard bottom-up proteomics protocol which included lysis using 1 % SDS buffer, reduction (DTT), alkylation (IAD), overnight acetone precipitation, and digestion using trypsin in 1:50 (enzyme to sample) ratio for 12 hours at 37 °C.[71] The protein digests were dried

using a vacuum concentrator (Centrivap, Labconco) and reconstituted in 1% (v/v) formic acid (FA) in LC-MS grade water.

Mass Spectrometry. For method development: About 0.5 µg of total tagged peptides were separated on a nanoflow LC system (Dionex Ultimate 3000 RSLC, Thermo). The peptides were separated on a 200 cm micropillar array column (µPAC, Thermo) in line with a C18 precolumn (100 µm inner diameter, 5 µm particle size with 100 Å pores, 2 cm length, Acclaim PepMap 100, Thermo). Peptides were trapped on the precolumn in Buffer A (LC-MS grade water, 0.1% formic). at 5 µL/min for 5 min and separated using a 200 min stepped gradient ramping Buffer B (LC-MS grade ACN, 0.1% formic acid) at 600 nL/min as follows: 1% for 0–5 min, to 7% in 15 min, to 25% in 115 min, to 32% in 25 min, to 45% in 33 min, to 75% in 7 min, held at 75% for 8 min, then ramped down to 2% in 2 min and equilibrated at 2% for 30 min. The nanoESI spray was generated by applying a positive voltage to a stainless steel emitter (Thermo). Ionized peptides were detected on a quadrupole-ion trap-orbitrap tribrid ultrahigh-resolution mass spectrometer (Orbitrap Fusion Lumos, Thermo) using the following parameters: survey (MS^1) at 120,000 FWMH resolution in the orbitrap analyzer; scan range, 300–1500 Da; Max IT, 50 ms; AGC target, 4×10^5 counts; microscans, 1; monoisotopic peak determination, peptide; dynamic exclusion, 60 s. Using data dependent acquisition executing 3 s cycle time, the precursor ions were isolated with a 1.6 Da window (±0.8 Da) and fragmented with high energy collision-induced dissociation (HCD) using nitrogen at 30% normalized collision energy with an intensity threshold of 5×10^3 counts The tandem mass spectra were acquired in the ion trap (AGC target, 1×10^4 counts; max IT, set to dynamic mode; microscans, 1).

Data Analysis. Raw HRMS were processed in Proteome Discoverer (version 2.2, Thermo Scientific) with SEQUEST searching against the *Xenopus laevis* proteome (version 9.2)

downloaded from Xenbase (46,000 entries) and mRNA derived PHROG database[183]. Search parameters included trypsin as the enzyme, allowed up to 2 missed cleavages, fixed modifications used were cysteine carbamidomethylation, methionine oxidation was set as a dynamic modification, all searches included a minimum of 1 unique peptide, 10 ppm mass tolerance for precursor masses, 0.6 Da mass tolerance for MS^2 fragment masses. Peptides and proteins were filtered with <1% false discovery rate (FDR). Unique and razor peptides were used for quantification. Proteins were quantified based on their label-free quantification. Data filtering and normalization were performed in MetaboAnalyst 3.0[203] or Perseus[204]. GO was annotated in the PANTHER classification system (version 16.0).

Safety. Personal protective equipment was used to handle chemicals, biological samples, and sharp items. Special care was taken in the handling of forceps, needles, and nanoESI emitters, all of which present a potential needlestick hazard.

Scientific Rigor. For method development, 5–6 BRs were prepared and processed to test each condition. Each sample was analyzed three times (three technical replicates). Biological replicates consisting of single dissections were collected over several batches of embryos on different days to account for biological variability.

3.4 Results and Discussion

Development of the µYODE approach for limited cell populations. Several studies have shown how highly abundant proteins interfere with the detection of proteins in low abundance during MS analysis across various species.[163, 168, 205] Particularly, in our previously published study we demonstrated how this interference is exaggerated in amphibian species due to yolk proteins. The total protein abundance in developing *Xenopus* embryos is predominantly

represented by the four vitellogenin or yolk proteins (Vtg), accounting for over 80–90% of the composition.[196] Yolk proteins in *Xenopus* are present in organelles called yolk platelets which can be separated physically. Tissues are lysed in yolk depletion buffer containing sucrose and centrifuged to slowly pellet the yolk platelets at the bottom of the vial. The supernatant devoid of yolk platelets is then collected and processed further using standard proteomics steps. However, such physical separation requires a large pool of tissues and proves difficult to downscale for limited tissue amounts.

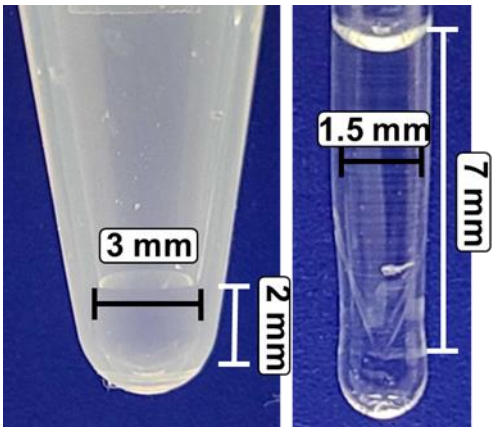

Figure 3.1. Demonstration of microscale yolk depletion idea. The difference in height of a limited volume of lysate (10 µL) deposited in a vial versus a capillary.

Previously we developed a yolk depleted carrier strategy to preclude the need for individual yolk depletion of the amount limited tissues (~200 ng of yolk free protein). In developing embryos, the composition of the carrier needs to be carefully evaluated to match the biological tissue under study. To further supplement the YODEC method we envisioned the idea of microscale yolk depletion (µYODE) to separate yolk platelets from small tissue amounts or sorted populations. As demonstrated above in **Figure 3.1** and earlier in Chapter 2 (**Figure 2.1)**, when processing small amounts in a standard vial, the main challenge lies in effectively aspirating yolk depleted supernatant from the pellet. Due to limited volume, the pellet is often disturbed and mixed with

the supernatant when aspirating with a pipette. To circumvent this challenge, we proposed performing yolk depletion in a custom capillary tube rather than a standard vial. **Figure 3.1** shows 10 µL of buffer deposited in a vial vs capillary. Due to the narrow diameter of the capillary, the height of the same volume of liquid in the capillary is 7 mm as opposed to 3 mm in the vial. Such vertical augmentation provides a clear distinction between the supernatant and the yolk pellet permitting easy contamination-free collection of the depleted supernatant.

Designing the microscale setup. The first step of designing the approach involved deciding the diameter of the capillary to allow deposition and retrieval of the sample lysate. We chose capillaries with varying inner diameters which were sealed on one end using a burner to create a 'capillary vial'. We found that depositing samples in diameters below 1.5 mm was challenging even with narrow taper pipette tips such as gel loading pipette tips (**Table 3.1**).

Table 3.1. Capillary selection for microscale setup.

Capillary dimensions (I.D. in mm)	Sample Deposition	Yolk pellet visualization	Extracting supernatant
1.1	Could not deposit, ID too narrow	-	-
1.17	Yes, with gel loading pipette tips	Yes	No, retrieval difficult due to losses
1.56	Yes, with gel loading pipette tips	Yes	Yes

To ensure complete aspiration of the supernatant, careful consideration was given to optimizing both the capillary sealing conditions and retrieval methods. **Figure 3.2A–3.2C** shows several conditions tested to ensure complete retrieval of supernatant without sample losses. To improve the distinction between yolk platelet and supernatant we tried using glycerol to create a density gradient or coating the bottom of the capillary with agar, however, both these methods did not allow clear separation. Alternatively, we observed that capillaries sealed to create a bulbous

round end facilitated robust and reproducible aspiration compared to capillaries that were pulled to create a conical end (**Fig. 3.2C**). To deposit and retrieve samples, we opted for narrow taper pipette tips initially due to ease of use, however, we observed inconsistent results when retrieving the supernatant. The yolk pellet occasionally aspirated along with the supernatant or losses were observed due to the long taper of the tip. To circumvent this challenge, we pulled thin micropipettes from 0.75 mm borosilicate capillaries with long taper (**Fig. 3.2D**). The borosilicate capillaries were connected to a syringe pump to aspirate the supernatant.

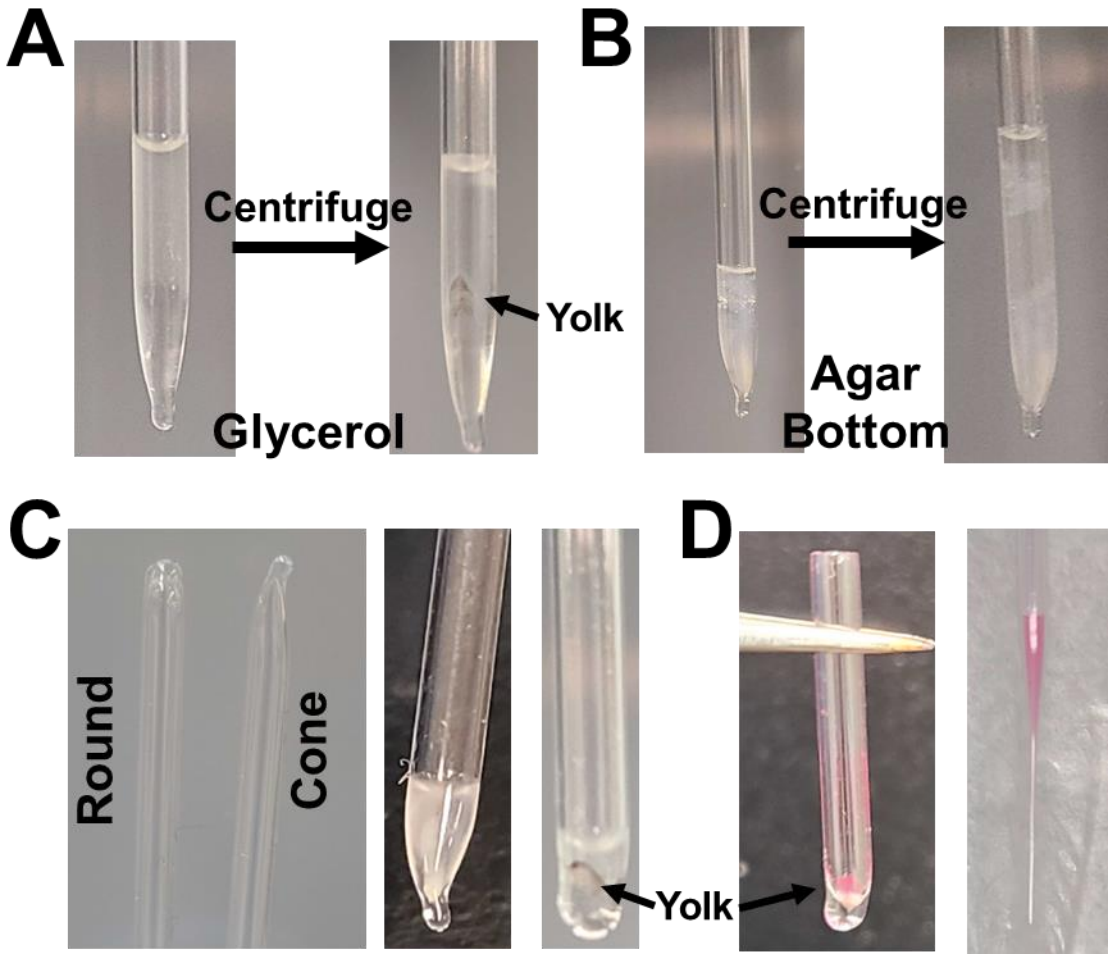

Figure 3.2. Design optimization for µYODE **(A)** Glycerol gradient to separate the yolk pellet and supernatant. However, the mixture homogenizes after centrifugation leaving no clear boundary. **(B)** Conical bottom was sealed with 2% (w/v) agar to reduce the loss of lysate in the

conical bottom. Visualization of yolk pellet with agar is challenging. **(C)** Distinction between cone bottom and rounded bottom capillary. A rounded bottom facilitates supernatant retrieval compared to a conical bottom. **(D)** Capillary micropipette to retrieve supernatant carefully without disturbing the yolk pellet. Pink dye is added to aid visualization.

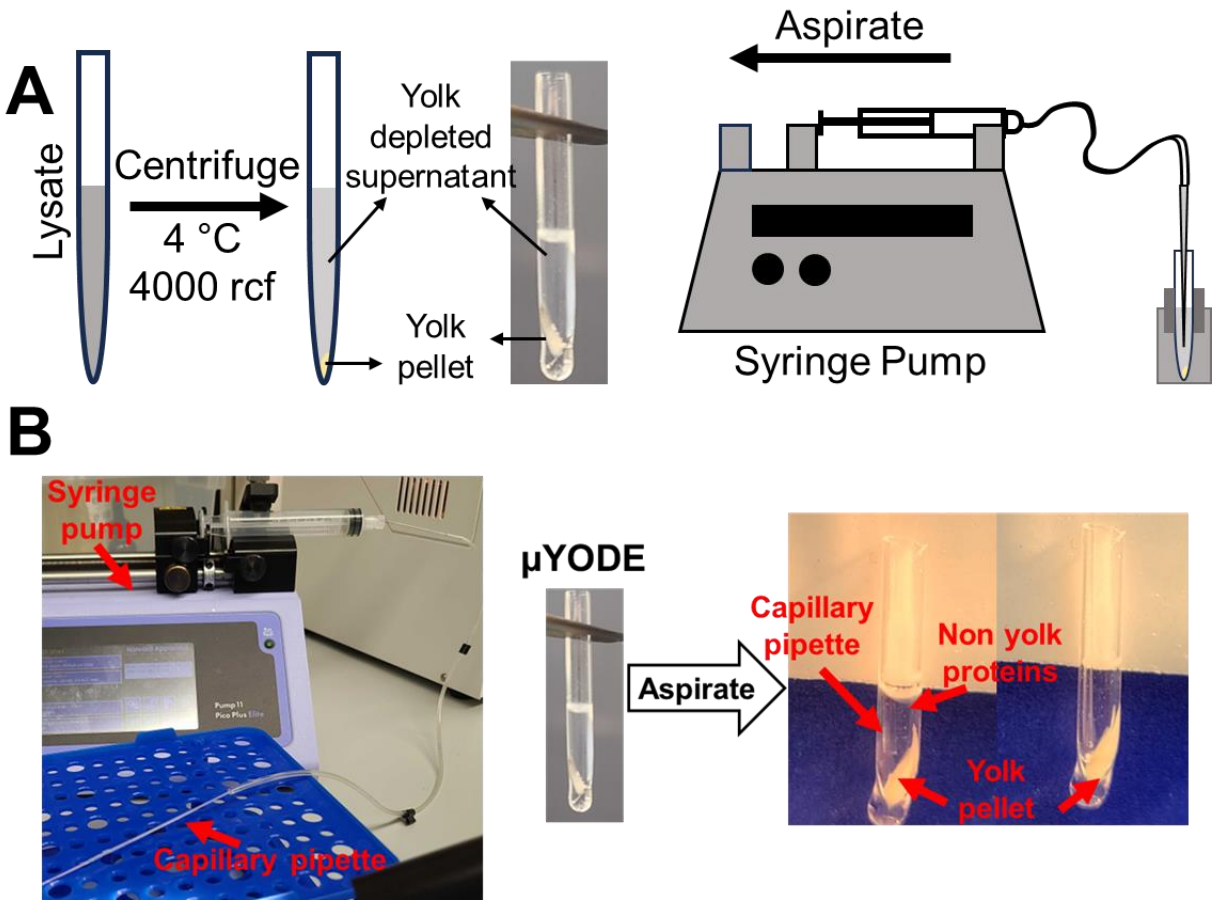

Figure 3.3. µYODE set up. **(A)** Schematic of working principle of µYODE **(B)** Demonstration of microscale depletion of neural cell clones

Figures 3.3A and 3.3B show the schematic of the µYODE approach, the microscale yolk depletion was carried out by depositing sample lysate in the capillary, the capillary was then placed in a holder and centrifuged to pellet the yolk platelets (*see Methods*). The capillary was then mounted vertically in a stable position to allow clear visualization of the pellet and withdraw the supernatant. The supernatant was carefully aspirated in a controlled manner using a micropipette connected to a syringe pump. The supernatant was then deposited in a low volume

lo-bind tube by applying forward pressure and processed further using bottom-up proteomic steps (*see Methods*).

Method validation. We traced and dissected the lineage of D11 cells (500–1000 cells) from a single embryo at the beginning of gastrulation (NF Stage 10). We chose the D11 lineage as it encompasses parts of both the SMO and the NE lineage which were the lineages of interest in the spatial study and for the ease of collection of large number replicates for method development. We compared the µYODE approach with three other control conditions to evaluate its performance which included conventional in-vial yolk depletion to test if µYODE is more consistent than the in-vial yolk depletion for limited volumes, standard in-vial bottom-up proteomics (without yolk depletion), and standard bottom-up proteomics in a capillary where all steps were carried out in the glass capillary. The standard in vial bottom-up proteomics allowed assessment of sample losses due to additional yolk depletion steps for small cell populations whereas in-capillary condition was added to assess sample losses due to adsorption on glass surfaces. Protein digests were then analyzed using LC-MS and quantified by label-free quantification.

Figure 3.4A displays the comparison of percent yolk PSMs and percent yolk abundance based on label-free quantification. As expected, the number of yolk PSMs and the relative yolk abundance in samples processed using standard bottom-up proteomics was found to be very high (70–80% of the total abundance). Although yolk depletion in a vial reduced yolk abundance compared to standard samples, the large error bars indicate it did not reproducibly deplete samples across BRs. In contrast, the µYODE approach performed consistently to deplete the yolk from limited cell populations and volume across all 6 BRs. **Figure 3.4B** displays the protein identifications from six BRs and three technical replicates between the µYODE and the control

conditions. The microscale approach achieved the highest number of protein IDs whereas the dispersed data points for other control conditions further corroborates that µYODE can consistently deplete yolk platelets and provide reproducible improvement in proteomic coverage. Over ~1300 proteins were exclusively identified by µYODE compared to conventional yolk depletion in a vial from single dissections of neural tissues. Histogram distribution of label-free quantification abundances revealed that the proteins exclusively identified using µYODE belonged to a lower abundance range than proteins commonly identified between both experiments (**Fig. 3.4C**).

Figure 3.4D shows the GO annotations for molecular functions for proteins exclusively detected using the µYODE approach. Newly identified proteins include several biologically relevant proteins such as transcriptional and translational regulators, molecular adaptors, and enzymes driving important catalytic reactions. For example, transcription factor Sp6 (Sp6), activating transcription factor 7-interacting protein 1 (Atf7ip), homeobox protein Tgif2 (Tgif2), POU domain, class 5, transcription factor 1 (Pou5f1), Myelin regulatory factor (Myrf), mothers against decapentaplegic homolog 1 (Smad1), and metastasis-associated protein Mta2 (Mta2) with known roles in cell cycle regulation, neural differentiation, cell proliferation, migration, and maintenance of pluripotency were detected exclusively in the µYODE dataset.[206-211] Furthermore, the microscale approach enabled deeper coverage of signaling pathways such as Wnt signaling, integrin signaling, cholecystokinin receptor (CCKR) signaling, fibroblast growth factor (FGF) signaling, and FAS pathway which are known to guide embryonic cell differentiation and tissue induction during development.

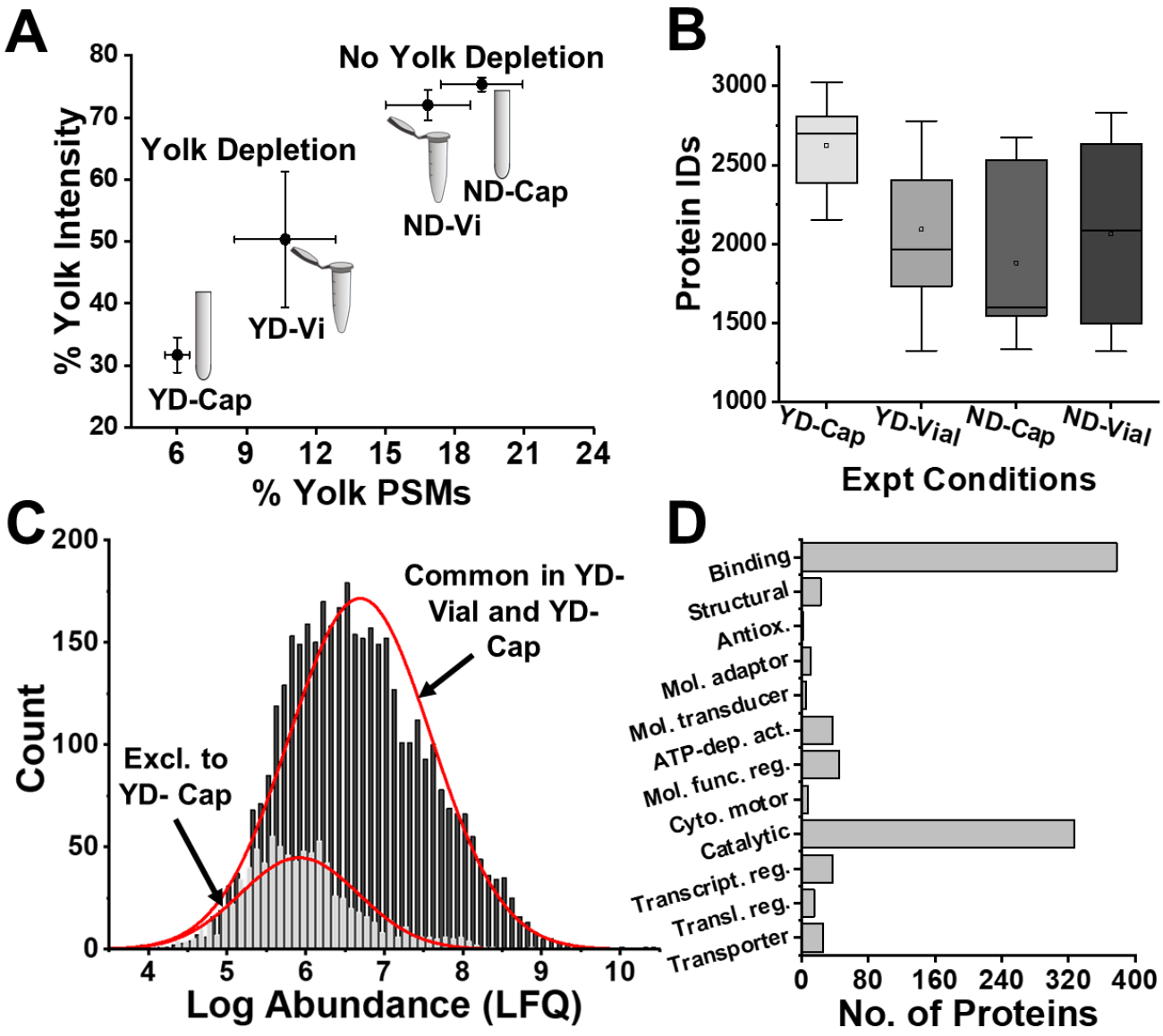

Figure 3.4. Method validation of µYODE approach. **(A)** Comparison of % yolk abundance and % yolk PSMs between µYODE (YD-Cap), conventional yolk depletion in a vial (YD-vial), standard bottom-up proteomics without yolk depletion in a vial (ND-vial) or a capillary tube (ND-Cap). **(B)** Protein identifications across 6 biological replicates and 3 technical replicates across different sample processing conditions. **(C)** Distribution of commonly quantified between µYODE and conventional in-vial yolk depletion and proteins exclusively quantified by µYODE using label-free quantification. **(D)** GO annotations for molecular function of proteins exclusively detected using the µYODE approach.

Enhancing coverage using new-generation LC separation. To further deepen the coverage from embryonic cell populations we sought to reduce the sample complexity before MS analysis. We employed the micropillar array based (µPAC) LC separation to improve the separation efficiency and maximize the duty cycle. **Figure 3.5A** demonstrates the differences between traditional packed bed columns and µPAC. The broad distribution of particle size and stochastic distribution of the packing material in traditional packed bed RP columns pose the problem of peak broadening and reduced resolution. Thus, to reduce the eddy diffusion and increase the theoretical plates (based on the Van Deemter equation), we chose to evaluate µPAC separation technology. The stationary phase is designed to reproduce a geometrical pattern of RP silicon wafers. Such an ordered array reduces dispersion of the molecules resulting in sharper peaks, higher resolution, and improved sensitivity. We performed extensive method development using commercial HeLa digest to test different parameters such as gradients, flow rate, and temperature to maximize protein identifications. We chose commercial HeLa digest to reduce variation due to sample preparation steps. As shown in **Figure 3.5A**, we achieved 2- fold higher proteome coverage with µPAC technology increasing the identifications from 1600 to 3600. We compared the peak widths and peak intensities of commonly detected precursor ions between packed bed and µPAC. As expected, the ordered geometry of µPAC improved resolution by narrowing peak widths which in turn resulted in sharper peaks and higher peak intensities (see *Appendix*, **Fig. A3**). We also found the newly identified proteins belonged to the lower abundance range compared to proteins identified by packed bed separation. (see *Appendix*, **Fig. A3**). Along with separation we also tested various nanoESI interfaces to improve ionization and detection sensitivity (see *Appendix*, **Fig. A4**).

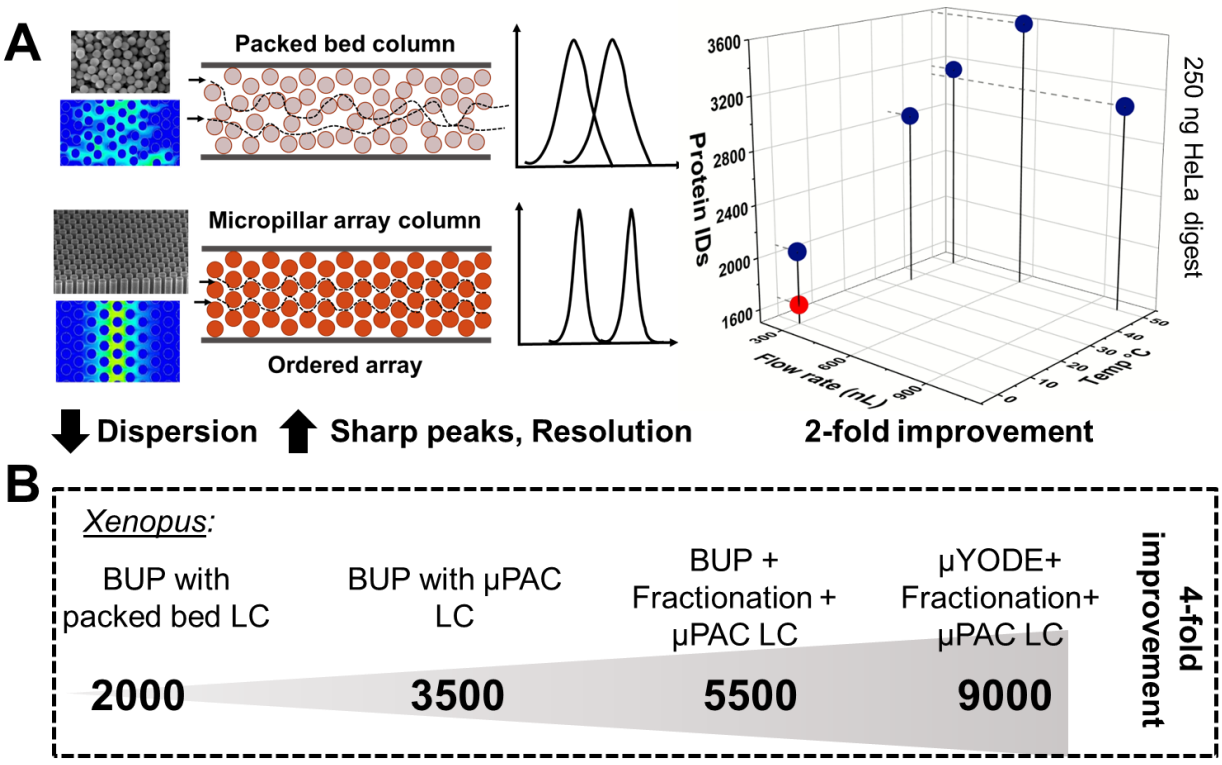

Figure 3.5. Deepening proteomic coverage using new generation LC separation. **(A)** Demonstration of working principle of packed bed vs µPAC separation. Optimization of flow rate, temperature, and gradient to maximize protein identifications using HeLa digest shows a doubling of protein identifications between packed bed versus µPAC. **(B)**. Deepening proteomic coverage in *Xenopus* tissues combining various separation modalities.

We observed similar improvements for *Xenopus* tissues with µPAC, to further resolve the complexity of the protein digest, we performed offline high pH fractionation prior to LC-MS. High pH fractionation prior to low pH LC-MS provides orthogonal separation as peptides elute based on their isoelectric point in different pH buffers. Furthermore, dividing the sample into multiple fractions reduces the dynamic range within each fraction, allowing for the detection of low-abundance proteins that might be masked by highly abundant ones in a non-fractionated sample. To enable offline fractionation, we combined µYODE with isobaric multiplexing which allows tagging different biological samples and combining them into one sample providing enough starting material for fractionation. **Figure 3.5B** displays improvement in proteomic coverage for *Xenopus* tissues as sample complexity is resolved by combining various analytical

modalities. Overall, the integration of µYODE with high pH fractionation and LC-MS analysis using µPAC yielded over 4- fold higher coverage compared to standard bottom-up proteomics.

3.5 Conclusions

Understanding mechanisms involved in various biological processes to understand healthy development and diseased states requires holistic knowledge of genes and gene products. Nevertheless, achieving deep proteomic coverage in global analysis remains a formidable challenge due to various factors pertaining to the large dynamic concentration range of proteins in the sample, sample complexity, and processing losses. In this work, we developed a microscale yolk depletion device, termed µYODE, to isolate yolk platelets from low volumes of tissue lysates while minimizing protein losses and eliminating the need for pooling large tissue amounts. The aforementioned strategy was developed as an alternative to the macroscale yolk depletion approach. It addresses the limitations in processing limited cell populations or volumes and demonstrates consistent and reproducible yolk depletion across multiple biological replicates from cell populations equivalent to ~500–1000 cells, particularly identifying lower abundance proteins critical for understanding cellular processes during embryonic development. µYODE can be used alone or in conjunction with the YODEC strategy (developed in the earlier chapter) to tailor the analytical method to the developmental stage of the embryonic tissues, protein amount available, and biological question under study.

Furthermore, we demonstrated that the integration of high pH fractionation and isobaric multiplexing with µYODE and µPAC LC-MS analysis quadrupled the proteomic coverage, highlighting the effectiveness of this comprehensive approach in overcoming challenges associated with sample complexity and limited sample availability. The enhanced sensitivity

offered sufficient improvement for a discovery based proteomic analysis to probe biological questions pertaining to tissue differentiation in the *Xenopus* model. Overall, the development and implementation of the µYODE approach for *Xenopus* tissues, in conjunction with advanced LC-MS technologies, represent a significant advancement in proteomic methodologies, offering researchers a powerful tool for in-depth analysis of complex biological samples and facilitating the discovery of novel proteins and pathways relevant to developmental biology and disease mechanisms.

Acknowledgements. Parts of this work were funded by the National Institute of General Medical Sciences (award 1R35GM124755 to P.N.), the University of Maryland-National Cancer Institute Partnership Program (seed award to P.N.), and the COSMOS Club Foundation (fellowship award to L.R.P.).

Chapter 4: Micro Proteomics for Uncovering Spatial Differences during Gastrulation

This chapter is based on the Pade, L.R. and Nemes, P. (2024) manuscript draft in preparation.

Contribution: L.R.P. collected and prepared the samples for proteomic analysis, analyzed the samples via LC-MS, performed data analysis, interpreted the results, drafted the manuscript, and revised the contents.

4.1 Abstract

Technologies enabling the identification of diverse types of proteins can improve our existing knowledge of mechanisms underlying tissue patterning and normal development. We studied two neighboring regions in an embryo, the SMO, which is an important signaling center, and the NE, which is induced by the SMO during gastrulation. The tissues later form mesodermal and neural tissues respectively. To understand proteomic differences involved in establishing spatially distinct tissue lineages, we performed deep proteomics on these cell lineages and compared them to the whole embryo. To maximize our protein identifications from limited tissue amounts (~200 ng yolk free protein in dissected clones from a single embryo), we used our previously developed analytical workflow which includes combining µYODE device with high-pH fractionation, followed by low-pH µPAC LC-MS to process and analyze the embryonic tissues. This integrated approach enabled ~8,000 protein identifications from lineage specific cell populations, providing 4-fold deeper coverage compared to our previous studies in pooled tissues. A comparison of the proteomic profiles of the two types of tissues and the whole embryo showed distinct enrichment in energy mechanism. For example, we found the enzymes involved in glycolysis and pyruvate metabolism pathways upregulated in mesodermal-fated organizer tissue, whereas citric acid cycle and fatty acid oxidation enzymes were enriched in the neural-

fated tissue. Furthermore, we uncovered the pivotal role the electron transport chain (ETC) plays in the organizer signaling. These encouraging results have led us to develop orthogonal functional assays to assess the role of energy cycles in tissue induction and differentiation enabling the discovery of important protein expression gradients crucial for normal tissue induction and formation during embryonic development.

4.2 Introduction

Vertebrate embryonic patterning is a topic of considerable research interest to understand healthy and diseased states. Spatial and temporal gradients of biomolecules during early embryogenesis guide key developmental events such as gastrulation resulting in tissue induction and formation of the three germ layers.[212] Specific groups of embryonic cells form transient vertebrate signaling centers active during gastrulation and play a critical role in embryonic development, especially in coordinating tissue induction and differentiation by sending signals to the neighboring cells. For example, the node in the human, chick, and mouse models and the Shield in the zebrafish model are known for their inductive signaling to guide gastrulation.[213] In mammals, the node is known to have a mesodermal fate but plays an important role in establishing bilateral symmetry within the embryo whereas xenotransplantation to other species such as chick displayed its ability to induce anterior structures.[214]

Similarly in the amphibian model, the SMO tissue is known to induce neural tissue from the ectoderm and guide dorso-anterior patterning.[1, 215] Initial embryology experiments in the early 1900s revealed the role of the organizer in inducing a secondary axis within a developing embryo. When the region of dorsal blastopore lip (organizer tissue) from one embryo was transplanted into another embryo, it prompted the formation of a secondary head and axis

primarily originating from host embryonic cells whereas the graft tissue gave rise to mesodermal and notochord tissue.[216] Such inductive signaling of the SMO holds the potential to answer biological questions related to neural tissue induction and formation which can be extrapolated to other models. In the following years, screening of genes and transcripts revealed important factors such as chordin, noggin, and follistatin are released from the organizer and signaling pathways such as Wnt/β Catenin, BMP signaling, and FGF signaling were important to pattern the embryo.[215, 217-218] Although several studies have uncovered important signaling mechanisms for gastrulation, our current knowledge is limited to a few genes and transcripts. To further expand our understanding of basic developmental biology related to the organizer function and its role in inducing neighboring cells, it is important to explore alternate directions that can provide complementary information to the genomic data.

Biological MS now offers a feasible solution to detect the translation of a wide array of proteins and metabolites simultaneously in a wide variety of biological specimens such as tissues to cells to organelles. Additionally, established cell fate maps in *Xenopus* offer a unique opportunity to trace the lineage of the organizer and the induced neural ectoderm within the developing embryo. Here, we harnessed the advantages of both the analytical technique and the biological model to probe into the proteomic gradients that help establish critical tissue lineages within the embryo and uncover noncanonical molecules involved in organizer signaling.

The goal of this study was to uncover spatial differences in the proteomic expression of two neighboring tissues known to develop into different tissue lineages. To facilitate deep proteome coverage, we used our microscale workflow developed previously (μYODE) without the need for pooling large dissection of tissues. Our μYODE approach enabled us to compare the

proteomes of the SMO, the NE which is induced by the SMO, and the whole embryo (WE) at the beginning of gastrulation. We uncovered differential enrichment in metabolic pathways between the SMO and the NE such as glycolysis, citric acid cycle, and electron transport chain. We found a gradient of reactive oxygen species (ROS) related metabolites exhibiting the highest enrichment in the SMO, a comparatively lower level in the NE, and the lowest concentration in the WE. We validated our proteomic findings with imaging experiments. Based on our proteomic data, we further designed experiments to perturb the metabolic enrichment and assess its effect on neural induction. Exploring the role of previously undiscovered molecules underscores the significance of reexamining crucial developmental events using a multi-omics approach, providing a comprehensive perspective on the molecular processes guiding embryonic development.

4.3 Experimental section

Materials and Reagents. All solvents for mass spectrometry were LC-MS grade. Solvents and reagent-grade chemicals were procured from Fisher Scientific (Pittsburgh, PA) or Millipore Sigma (St. Louis, MO) unless explicitly stated otherwise. Trypsin protease (MS grade), the TMT10plex isobaric label reagent set, and green fluorescent dextran (FD, Alexa Fluor 488, 3000 MW) were acquired from Fisher Scientific. Sodium dodecyl sulfate (SDS) and combretastatin A-4, cytochalasin D, nonidet P-40 substitute (NP-40), dithiothreitol (DTT), and iodoacetamide (IAD) were sourced from Sigma-Aldrich (St. Louis, MO). Borosilicate capillaries were obtained from Sutter Instrument (Novato, CA).

Solutions. Embryological solutions such as 100% (w/v), 50%, 20% Steinberg's solutions (SS), and 3% (w/v) ficol in 100% (w/v) SS were prepared according to established protocols for

embryo culture and tissue dissection.[201] Components of the yolk depletion buffer were the same as listed in Chapter 3 and its preparation is described in the following publications.[159, 183] Fixative for embryos and larvae contained 4% (w/v) paraformaldehyde in 3% (w/v) sucrose solution.

Animal Care and Embryology. Adult male and female *X. laevis* frogs purchased from Nasco (Fort Atkinson, WI) or Xenopus1 (Dexter, MI) were humanely maintained and handled in a breeding colony approved by the Institutional Animal Care and Use Committee of the University of Maryland, College Park (R-FEB-21-07). Embryos were obtained via natural mating following previously published protocols.[201] After obtaining the embryos, the jelly coat was dissolved using 2% (w/v) cysteine solution (pH 8) to facilitate experimental manipulation and embryos were transferred to a sterile 100% SS solution. Cell fates in *Xenopus* are established based on the stereotypical pigmentation, thus stereotypical embryos were identified and sorted at the 2-celled stage and cultured to the 32-cell stage (Nieuwkoop and Faber, NF stage 5 or 6) in 100% SS.[8-9, 202] For the spatial study, cell clones of the NE tissue and the SMO tissue were traced by injecting green FD in their respective precursor cells identified based on their pigmentation and location (D111 for NE, D112, and D212 for SMO) at the 32-cell stage (NF stage 6). Incorrectly or damaged embryos were removed. The embryos were cultured to NF Stage 10 (beginning of gastrulation) in 50% SS in 14 °C incubator and monitored under a stereomicroscope using epifluorescence (model SMZ18 with GFP-B 480 filter, Nikon Instruments Inc., Melville, NY).

Tissue Collection. Cell clones arising from precursor cells of the SMO and the NE were manually dissected using sharp forceps in 50% SS in multiple replicates under the stereomicroscope. Single dissection of the SMO or the NE tissue obtained from one embryo was frozen in tubes immediately and stored at -80 °C until further processing.

Sample Preparation. For the spatial analysis, individual dissections of the SMO and the NE tissue were processed using the µYODE approach using 1.56 mm borosilicate capillary tubes while the whole embryos were yolk-depleted in standard Eppendorf tubes followed by standard bottom-up steps. Single dissections were lysed using 8 µL of yolk depletion buffer whereas 50 µL was used for the whole embryos. Yolk depleted supernatant was denatured using 1% SDS and processed further with standard bottom-up proteomics steps such as reduction, alkylation, and trypsin digestion. Six biological replicates of protein digests of the SMO, the NE tissues, and the whole embryo were barcoded using tandem mass tags (TMT 10plex). A pool channel containing a mixture of all tissues and biological replicates (BR) was included to assist data normalization. Samples were mixed in a 1:1 ratio based on peptide assay. A total of 6 BRs were prepared and divided into 2 sets of TMT experiments. Each TMT set contained 3 BRs of the SMO, the NE, the whole embryo, and one pool sample with randomized tag assignments. Tagged samples were then fractionated at high pH (pH= 10) using an RP column (InfinityLab Poroshell HPH-C18, Agilent) to collect 72 fractions (1260 Infinity II LC and fraction collector, Agilent) that were concatenated to 8 fractions, dried, and reconstituted in 0.1% FA in LC-MS grade water for further MS analysis. The concatenation strategy of 72 fractions to 8 fractions is described in the Appendix (**Fig. A5**)

Mass Spectrometry. The RPLC separation using a µPAC gradient remained consistent with the procedure outlined in Chapter 3. For spatial analysis: MS method was to enable synchronous precursor selection (SPS) MS^3 multiplexed quantification. The following parameters were used for TMT analysis: survey (MS^1) at 120,000 FWMH resolution in the orbitrap analyzer; scan range, 400–1600 Da; Max IT, 50 ms; AGC target, 4×10^5 counts; microscans, 1; monoisotopic peak determination, peptide; dynamic exclusion, 60 s. Using data dependent acquisition, the

precursor ions were isolated with a 0.7–Da window and fragmented with collision-induced dissociation (CID) using helium at 35% NCE with an intensity threshold of 5×10^3 counts. The tandem mass spectra were acquired in the ion trap (AGC target, 1×10^4 counts; max IT, set to auto mode; microscans, 1). The top 10 MS^2 fragments were isolated using synchronous precursor selection and dissociated (MS^3) with a scan range of 100–500 Da and fragmented with higher-energy collision-induced dissociation (HCD) in nitrogen at 65% NCE. The MS^3 spectra were acquired at 50,000 FWMH in the orbitrap (AGC target, 5×10^4 counts; max IT, auto mode; microscans, 1).

Data Analysis. Raw HRMS were processed in Proteome Discoverer (version 2.2, Thermo Scientific) with SEQUEST searching against the *Xenopus laevis* proteome (version 9.2) downloaded from Xenbase containing 46,000 entries and combined with mRNA derived PHROG database[183]. Search parameters included trypsin as the enzyme, allowed up to 2 missed cleavages, cysteine carbamidomethylation, and TMT label as fixed modification, methionine oxidation as a dynamic modification, all searches included a minimum of 1 unique peptide, 10 ppm mass tolerance for precursor masses, 0.6 Da mass tolerance for MS^2 fragment masses, and 20 ppm mass tolerance for TMT labels. Peptides and proteins were filtered with <1% false discovery rate (FDR). Unique and razor peptides were used for quantification. Proteins were quantified based on their MS^3 reporter ion abundances (for spatial analysis). Data filtering and normalization were performed in MetaboAnalyst 3.0[203] or Perseus[204]. For the spatial study, data was normalized using the pool sample to reduce batch effects and log-transformed for further evaluation. GO was annotated in the PANTHER classification system (version 16.0). Statistical overrepresentation test and fuzzy mean clustering were performed using PANTHER and VsClust respectively.[187] Volcano plots were constructed using Origin Pro software.

Metabolomics. For preliminary metabolomic analysis using gas chromatography MS (GC-MS), 20 dissections of the SMO and the NE tissue and 2 WE at Stage 10 were pooled for extraction. A mixture of isotopic standards of TCA cycle metabolites, pyruvate, and lactate were spiked in the samples in a known amount. Metabolites were extracted by adding chilled ACN to the respective tissues lysing using a bead mill. The lysate was centrifuged at 10,000 rcf to pellet the debris and the supernatant was separated and dried for further processing. Metabolites were converted to Tert-butyldimethylsilyl derivatives and analyzed on GC-MS using the previously published protocol. [219] GC-MS data was manually integrated to extract quantitative information.

H_2O_2 Live Labeling. Biotracker H_2O_2 live dye was reconstituted as per vendor protocol to prepare the 1 mM stock solution. The stock solution was diluted to 5 µM with 20% SS. Control or treated (H_2O_2 exposed**)** embryos were incubated in the dye for ~30 min followed by washing with 20% SS for 5–10 min. The wash step was repeated twice. Embryos were imaged using the stereomicroscope with a monochromatic camera (SMZ18, Nikon Instruments Inc) in the brightfield and GFP (488 nm) channels. Images were analyzed using the NIS Elements software (Nikon). The ratio of the fluorescence intensity in the SMO region and region below the blastopore lip was compared between conditions. Statistical differences were evaluated using a t-test or Mann Whitney U test.

H_2O_2 Live Culturing. H_2O_2 culture media was prepared by diluting 12% (w/v) H_2O_2 stock with 20% (v/v) SS solution to 100 mM and 50 mM solutions. Embryos before the beginning of gastrulation (NF Stage 8), at the beginning of gastrulation (NF St 10), and at post gastrulation (NF Stage 13 and NF Stage 17) were incubated in the H_2O_2 culture media for 2–15 minutes. Embryos were transferred to non-spiked culturing solution (20% SS) and allowed to develop to larval stages (NF Stage 32/34).

Fixing. Larvae from H_2O_2 live culturing experiments (Control or H_2O_2 exposed) were fixed using 4% (w/v) paraformaldehyde in 3% (w/v) sucrose solution by incubating for 1 hour on a nutator and washing with 1× PBS 3 times. The fixed larvae were stored at 4 ºC until further imaging.

Imaging and Scoring. Fixed larvae from H_2O_2 live culturing experiments (Control or H_2O_2 exposed) were imaged under brightfield using a stereomicroscope equipped with a color camera (SMZ18, Nikon Instruments Inc). To evaluate the effect of disturbance of the H_2O_2 gradient on the neural induction and dorsoanterior patterning, the diameter of the eye, length of the spine, and the area of the cement gland were measured for the control and experimental group. Statistical analysis (t. test or Mann Whitney U test) was performed to evaluate the significance of differences between control and experimental groups.

Safety. Personal protective equipment was used to handle chemicals, biological samples, and sharp items. Special care was taken in the handling of forceps, needles, and nanoESI emitters, all of which present a potential needlestick hazard.

Scientific Rigor. For method development, 5–6 BRs were prepared and processed to test each condition. Each sample was analyzed three times (three technical replicates). For the biological part of the study, samples were processed together to avoid variability in sample preparation. 6 BRs and 3 technical replicates were analyzed for the spatial study. TMT tag assignments were randomized between two TMT sets. A pooled sample prepared by making a mixture of all sample types and BRs was included in each TMT set which was used to normalize the data and reduce TMT batch effects. Imaging and culturing experiments were repeated over several batches of natural mating with different frog pairs to ensure statistical confidence.

4.4 Results and Discussion

In this study, we sought to compare the proteome of the SMO, a known signaling center, and induced neighboring tissue, the NE. Early embryonic experiments established that the SMO releases morphogens to guide the differentiation of neighboring cells to induce neural tissue and pattern the embryo. Although the SMO tissue and the induced NE tissue arise from similar dorsal cells at the early stage and lie adjacent to each other during gastrulation, they eventually give rise to two different lineages, mesodermal and neural respectively. How the proteomes of these neighboring tissues differ from each other remains to be understood. We performed a deep discovery proteomics study to uncover the spatial differences in protein translation between the SMO, the NE, and the whole embryo. Here, we utilized our previously developed µYODE approach to address the technical limitations associated with processing minuscule embryonic tissues.

Spatial proteomics approach. As displayed in **Figure 4.1** (top panel) to collect the SMO and the NE tissues, we lineage traced the cell clones by injecting fluorescent dye in the precursor cells. Based on the cell fate maps, we injected 1 nL of green fluorescent dextran (FD) in both left and right D112 and D212 cells (for SMO) and D111 cells (for NE) at the 32-cell stage (NF stage 6) using a microinjector. The injected embryos were cultured to the beginning of gastrulation which is marked by the appearance of the blastopore lip (NF stage 10). The descendent cells were dissected manually using sharp forceps under a stereomicroscope with epifluorescence. **Figure 4.1** shows SMO in green and NE in red dye to aid visualization of the neighboring tissues, however, all collections were performed by solely injecting green FD. A single dissection of the SMO or the NE from each embryo was frozen individually. Whole embryos at the same stage were also frozen individually. The single dissections were yolk depleted using the

µYODE approach while the whole embryos were yolk depleted in vials. A total of 6 BRs were processed for bottom-up proteomics. The protein digests were barcoded using TMT 10 plex tags for relative quantification and two sets of TMT batches were prepared to accommodate all 18 samples. Barcoded samples from each batch were combined and fractionated at high pH to reduce sample complexity. Each fraction was analyzed at low pH via LC-MS implementing SPS MS^3 method[220] for TMT based quantification (**Fig. 4.1**, bottom panel).

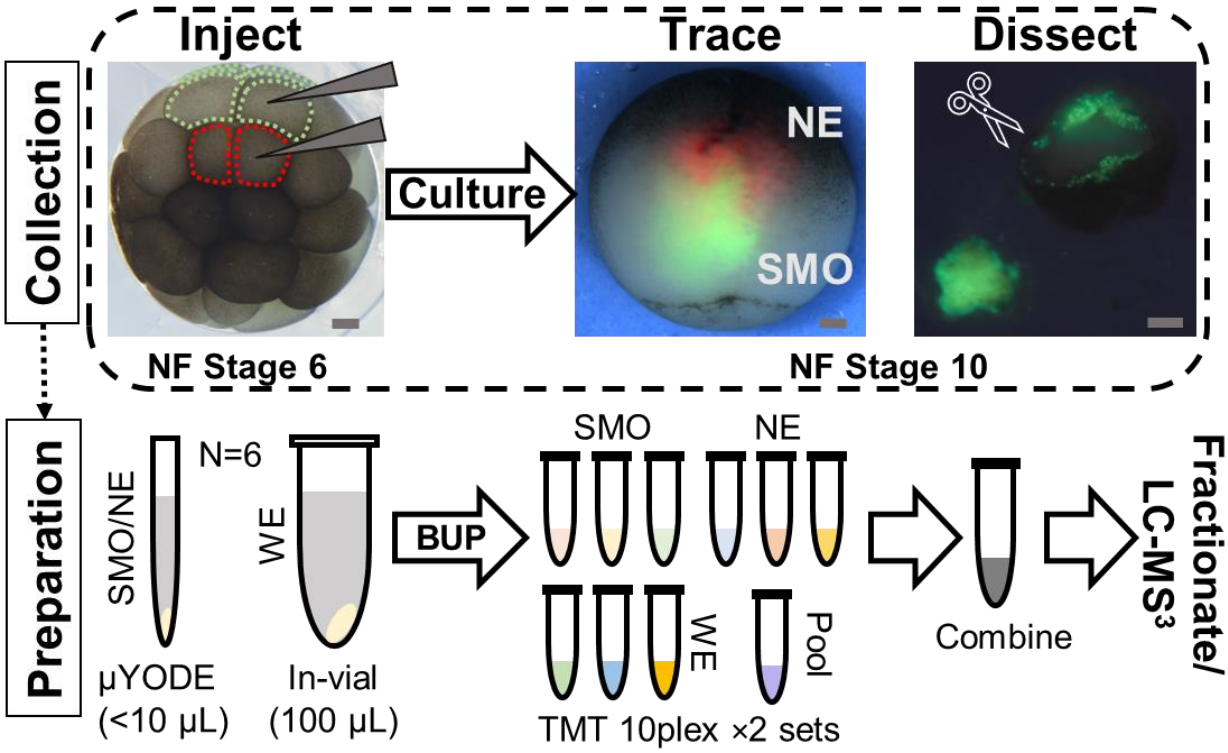

Figure 4.1. Strategy for spatial proteomics in gastrulating embryo. The SMO and the NE tissue were lineage traced by injecting FD in the left and right precursor cells at NF stage 6 and culturing the embryos to NF stage 10 for dissection (top panel). Individual single dissections (1 BR=1 dissection) of the SMO and the NE were yolk depleted using µYODE whereas single WE (1 BR=1 whole embryo) at the same stage were processed in vials. Samples were processed for bottom-up proteomics and barcoded with TMT tags. 2 10 plex sets were prepared with each set containing 3 BRs of the SMO, the NE, the whole embryo, and one pooled sample. Barcoded samples within each set were combined and fractionated at high pH and further analyzed using the LC tandem MS method (bottom panel). Key. BUP, bottom-up proteomics. Scale bar: 100 µm

Deep spatial proteomics of the gastrulating embryo. We identified ~8000 proteins cumulatively between two barcoded TMT sets containing six biological replicates of the SMO, the NE, and the WE tissue. We achieved one of the deepest coverages achieved from such a limited amount (1 single dissection ≤ 500 ng of yolk free protein amount) of tissues. Compared to our previously published studies in pooled tissue (~50 µg) samples from similar tissue lineages we more than doubled the coverage (**Fig. 4.2A**). Both of our previously published studies utilized pooled tissues (N=5 as 1 BR) with a total of three BRs, employing our µYODE approach we were able to preclude pooling and averaging signals over multiple pooled biological tissues while simultaneously increasing the number of BRs to facilitate rigorous statistical analysis. We used one single dissected clone as one BR (N=1 as 1 BR) and collected six individual BRs for each tissue with increased throughput. The Venn diagram (**Fig. 4.2A**) comparing the proteins identified using the microscale approach to the previously published study on the pooled SMO tissues shows over ~4700 proteins exclusively identified using µYODE. Further, in depth, GO analysis revealed a broader coverage of diverse proteins belonging to important signaling pathways pivotal for embryonic development such as FAS, Wnt, TGF-beta, Integrin, RAS, epidermal growth factor (EGF), ubiquitin proteasome pathways compared to the previous analysis on pooled tissues (**Fig. 4.2B**). In contrast to the ~450 proteins previously found to be differentially expressed between the SMO tissues and the whole embryo in the pooled dataset[221], the enhanced sensitivity enabled us to expand this coverage to ~2600 proteins for the same analysis from single clones. Overall, the dataset revealed ~3500 proteins to be differentially regulated in the SMO, the NE, and the whole embryo at the beginning of gastrulation. Such improvement in sensitivity paved the way for analyzing canonical and noncanonical differences between these tissue types to drive biological discovery.

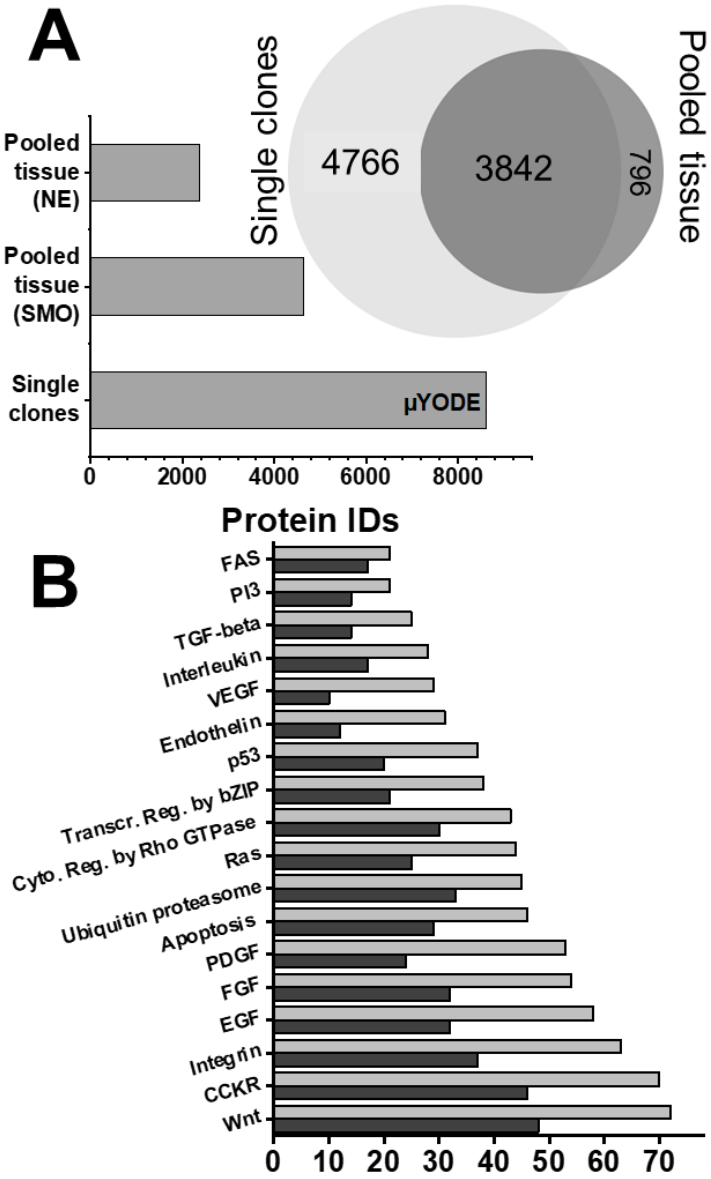

Figure 4.2. Comparison of deep proteomics in single dissected clones using μYODE approach to previously published studies in pooled tissues. **(A)** Proteins identified in pooled SMO or NE tissues via conventional in-vial yolk depletion versus single dissections of SMO or NE processed via the μYODE approach**.** Venn diagram comparing the overlap of pooled SMO dataset versus μYODE dataset. **(B)** GO annotation for signaling pathways covered under pooled dataset (dark gray) versus single clones (light gray) shows μYODE enabled a deeper coverage of proteins involved in different signaling pathways. Key: PI3, phosphoinositide 3-kinases; VEGF, vascular endothelial growth factor; Transcr. Reg., transcription regulation; Cyto. Reg., cytoskeletal regulation; PDGF, platelet-derived growth factor; FGF, fibroblast growth factors; EGF, epidermal growth factor; CCKR, gastrin, and cholecystokinin receptor.

Differential proteomic translation between neighboring tissues. At first, we sought to investigate if the three tissue types could be distinguished based on their proteome, therefore we performed unsupervised hierarchical-cluster-analysis. **Figure 4.3** shows the heatmap of 50 statistically different proteins between the SMO, the NE, and the WE. We were able to distinguish these tissues based on their proteome, confirming regional or spatial differences with the embryo. Surprisingly, despite the close spatial proximity between the SMO and the NE as they lie adjacent to each other, the proteomic profile of the NE is analogous to the WE than the SMO. The distinctive proteomic enrichment in the SMO could indicate its role as a transient signaling center. We annotated the statistically different proteins for protein functions and signaling pathways to unveil distinctively enriched categories (**Table A2**). The ontology revealed electron transport chain (ETC) related proteins Nudfs6 and Nudfs4 and cytoskeletal regulating protein Flna were enriched in the organizer. The enrichment of ETC related proteins in the SMO correlated well with our previous finding from the pooled organizer tissue.[221] We further performed STRING analysis to find common patterns in the proteins found enriched in the NE tissue (**Fig. 4.4**). The proteins enriched in the NE tissue were divided into three main categories which included proteins belonging to transcription and translation regulation, proteasome machinery, and enzymes related to carboxylic acid metabolism such as argininosuccinate lyase (Asl) and 3-hydroxyacyl-CoA dehydrogenase (Hadh). Evidence suggesting deficiency in Hadh and related enzymes during embryonic development leading to peripheral neuropathy and retinopathies in humans has been reported.[222] The active upregulation of proteasome and translation related proteins could suggest remodeling of the proteome to drive differentiation of ectoderm to neural ectoderm. Upon comparing the NE tissue enriched proteins identified in our heatmap analysis with genes enriched in neural precursor cells induced from human embryonic

stem cells[223], we observed an overlap between several proteins and the transcriptomic data such as Adss (adenylosuccinate synthase), Rps15 (ribosomal protein s15), Eif2s1 (eukaryotic translation initiation factor 2 subunit alpha), Capn2 (calpain-2), Ipo9 (importin 2), and Rps6ka3 (ribosomal S6 kinase).

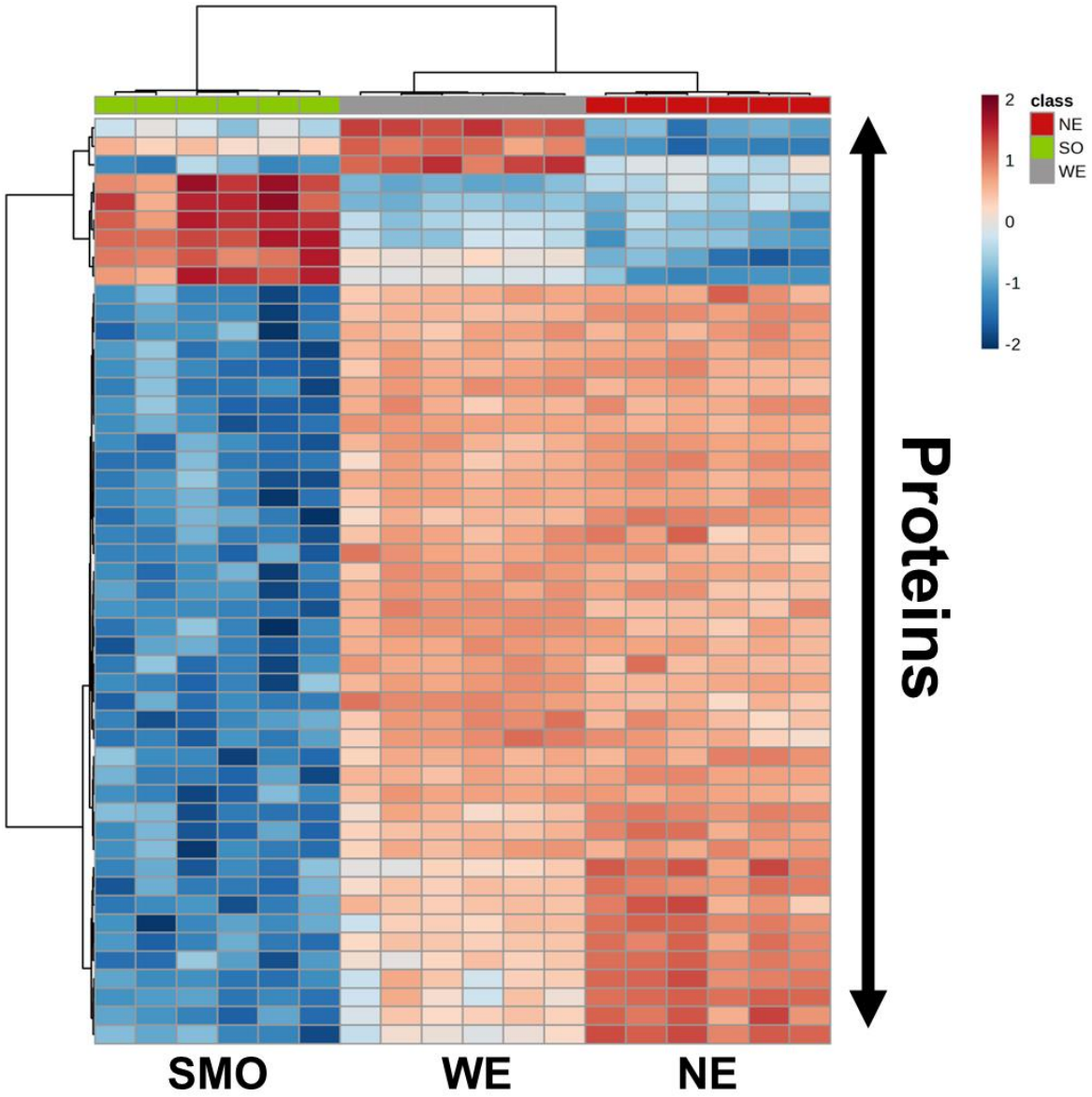

Figure 4.3. Unsupervised hierarchical clustering (heatmap) displaying 50 statistically significant proteins that showed distinctive enrichment in the SMO, the NE, and the WE. (Color bar represents fold change)

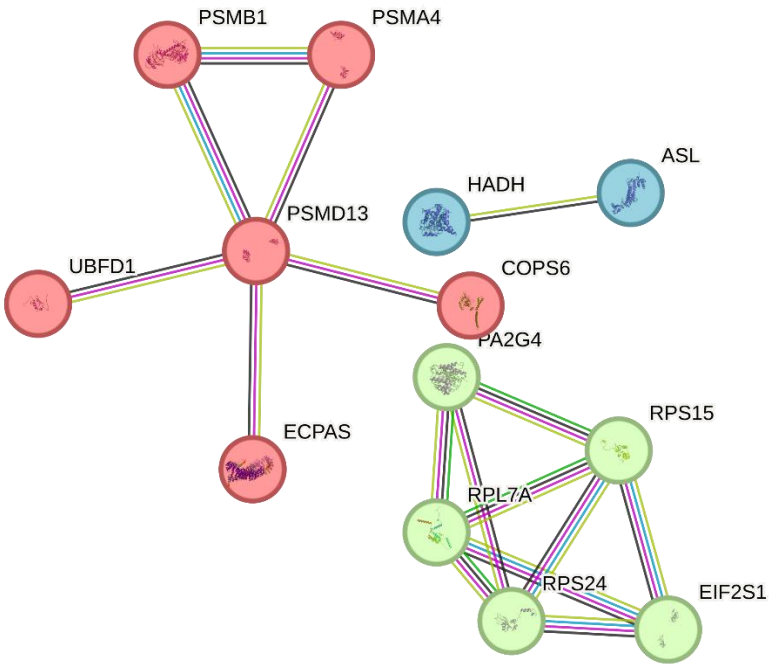

Figure 4.4. STRING analysis to find correlation between proteins found upregulated in the NE tissue.

Canonical differences between the SMO and the NE tissue. To uncover the spatial differences between the neighboring SMO and the NE tissues, we performed statistical analysis on quantified proteins to identify differentially enriched proteins. Initially, we compared the SMO and the NE proteome against the WE proteome using the volcano plot. Next, we conducted statistical overrepresentation tests on the proteins found to be statistically upregulated in the SMO and the NE tissue compared to the WE to uncover GO categories that are overrepresented in these tissue types. The overrepresentation test queries the list of proteins from the experimental results against the reference proteome to assess if a particular class (molecular function, biological process, signaling pathway) is overrepresented in the experimental dataset. We used the WE (at the same stage) as a reference since the whole embryo constitutes both these spatial regions and would provide an averaged or overall protein abundance at that stage. **Figure**

4.5 shows the volcano plots and the overrepresented signaling pathways for the proteins enriched in the SMO and the NE tissue.

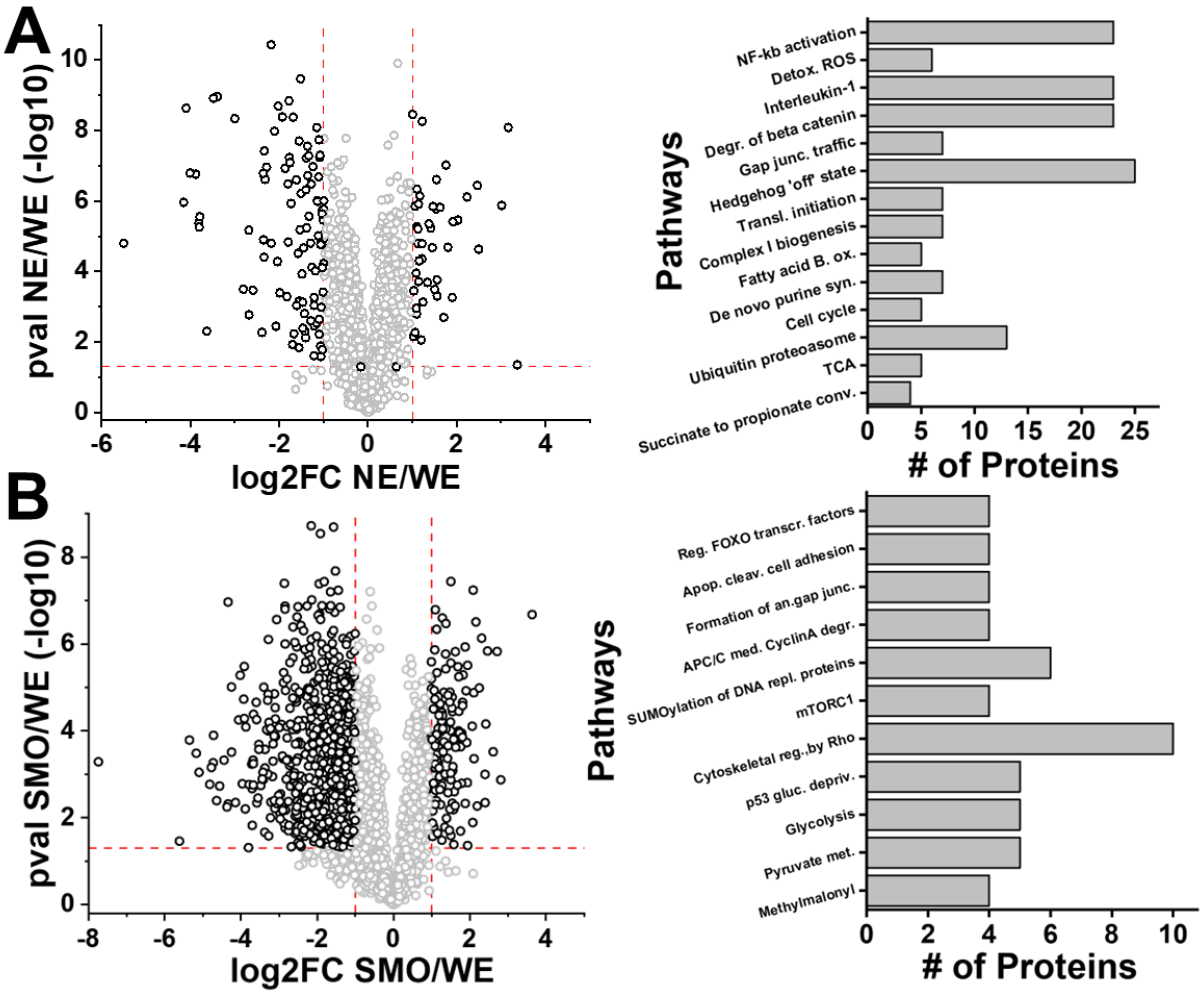

Figure 4.5. Volcano plot displaying statistically upregulated and downregulated proteins in the SMO vs WE and the NE vs WE conditions. **(A)** Statistically overrepresented pathways in the NE. **(B)** Statistically overrepresented pathways in the SMO. Key: NF-kb activation, FCERI mediated NF-kB activation; Detox. ROS, detoxification of reactive oxygen species; Interleukin-1, interleukin-1 signaling; Degr. of beta catenin, degradation of beta-catenin by the destruction complex; Gap junc. Traffic, gap junction trafficking and regulation; Hedgehog 'off' state, hedgehog signaling 'off' state; Transl. initiation, translational initiation; Fatty acid B. ox, fatty acid beta oxidation; De novo purine syn., De novo purine biosynthesis; Ubiquitin proteasome, ubiquitin proteasome pathway; TCA, tricarboxylic acid cycle; Succinate to propionate conv., succinate to propionate conversion; Reg. FOXO transcr. Factors, regulation of FOXO transcription factors; Apop. cleav. cell adhesion, apoptotic cleavage of cell adhesion proteins; Formation of an. gap junc., formation of annular gap junctions; APC/C med. CyclinA degr., Cdc20:Phospho-APC/C mediated degradation of Cyclin A; SUMOylation of DNA repl. proteins,

SUMOylation of DNA replication proteins; mTORC1, mTORC1 signaling; Cytoskeletal reg. by Rho, cytoskeletal regulation by Rho GTPase; p53 gluc. depriv., p53 pathway by glucose deprivation; Pyruvate met., pyruvate metabolism; Methylmalonyl, methylmalonyl pathway.

Figure 4.5A highlights differences in protein concentrations between the NE and the WE and the right panel showcases the GO signaling pathways upregulated and overrepresented in the NE tissue. Many of these signaling pathways/ signaling proteins can be traced back to their canonical roles in inducing neural tissue or differentiating embryonic stem cells to neural cells in the developing embryo across various models. For example, NF-kb transcription in stem cells has been shown to be pivotal in the differentiation of embryonic stem cells to neural progenitor cells.[224] Additionally, tight regulation of NF-kb activation is required to control mesodermal and neuroectodermal cell fates.[225] NF-kb activation in turn is regulated by ROS which is also seen in the proteomics dataset as we see the upregulation of antioxidant enzymes that control detoxification of ROS species.[226] Earlier transplantation and genomic experiments conducted in the *Xenopus* and Chick models report that blocking of Wnt/ β catenin signals promotes the induction of neural fate from ectodermal cells.[227] In concurrence, we found upregulation of the degradation of β catenin by the destruction complex. One of the key regulators of the destruction complex, a Serine/threonine-protein phosphatase namely Ppp2r5a[228] (or Pp2a) was found to be significantly upregulated in the NE tissue compared to the WE as well as the SMO further corroborating our proteomic data with the established literature. We also found several metabolic pathways such as beta fatty acid oxidation upregulated in the NE tissue, whose role has also been explored in human neural development.[222]

Figure 4.5B reveals the pathways enriched in the SMO tissue. Similar to the NE tissue, we conducted a deep literature review to uncover the known canonical roles of these pathways in previously published literature. We found several parallels between our proteomic dataset and

earlier genomic/transcriptomic studies. Regulation of FoxO (Forkhead box, class O) transcription factors is required for proper closure of the blastopore lip[229]. As the Spemann organizer is situated right above the blastopore lip (see **Figure 4.1**), the higher abundance of FoxO-regulating proteins in the SMO is substantiated. FoxO proteins are phosphorylated by Akt-1 which in turn controls its translation and was found significantly upregulated in the SMO tissue compared to the WE and the NE in our dataset. As cells start migrating and moving inwards over the blastopore lip during gastrulation, enrichment of apoptotic cleavage of cell adhesion proteins is logical to diminish cell to cell contact and allow controlled movement of cells. This process is further complemented by the upregulation of cytoskeletal regulation by the Rho GTPase pathway which modulates the organization of the actin cytoskeleton to control cell protrusion, motility, and adhesion. Signaling pathways including mTORC1 and p53 have known roles in eliminating defective pluripotent cells to reduce incorrect patterning of cells during differentiation,[230] or act as redox sensors, and help cell survival during oxidative stress.[231] We also found upregulation of SUMOylation which is an important modification to regulate subcellular localization of proteins and protein function. [232] Apart from canonical pathways, we also found certain higher activity of metabolic pathways such as glycolysis in the SMO tissue. Metabolic pathways such as glycolysis and OXPHOS play an important role in maintaining the stemness of embryonic cells and cell differentiation, hence holding the potential to discover unknown roles or metabolic enzymes and small molecules in controlling embryonic development.[233]

Non-canonical proteomic differences between the SMO and the NE tissue. Subsequently, we aimed to classify proteins according to their abundance profiles in order to discern proteomic patterns within the three tissue types. We utilized the fuzzy c-means soft clustering method to visualize the proteomic gradients within the organizer, induced NE, and WE. The clustering

analysis was conducted using the freely available online tool, VsClust. The proteomic profiles were grouped into 7 clusters (**Fig. 4.6**). We conducted statistical overrepresentation tests on each cluster to pinpoint alterations in the translation of proteins associated with specific GO categories. We annotated the dataset for pathways, biological processes, and molecular functions. Notably, we found several metabolic pathways differentially regulated between the SMO, the NE, and the whole embryo.

Although several studies discuss the pivotal role of metabolism in the context of stem cell differentiation, very few recent studies have focused on these lesser-known mechanisms in a three dimensional whole embryo model.[233-235] Even in the *Xenopus* model, a significant area of investigation focused on determining the metabolic activity within the organizer and investigating whether this activity played a role in defining the inductive signals responsible for patterning the embryo before the era of modern molecular genetics. Many of the early publications in this field highlighted oxygen consumption as a key metric for characterizing metabolic activity during gastrulation. However, due to a lack of sophisticated techniques and a shift towards genomic analysis, the significance of the metabolic profile of the organizer (SMO) remains ambiguous. The recent resurgence of interest in developmental metabolism, facilitated by new assays and advancements in MS, has provided an opportunity to reassess the role of metabolism in patterning and inducing cell fate at the organizer.

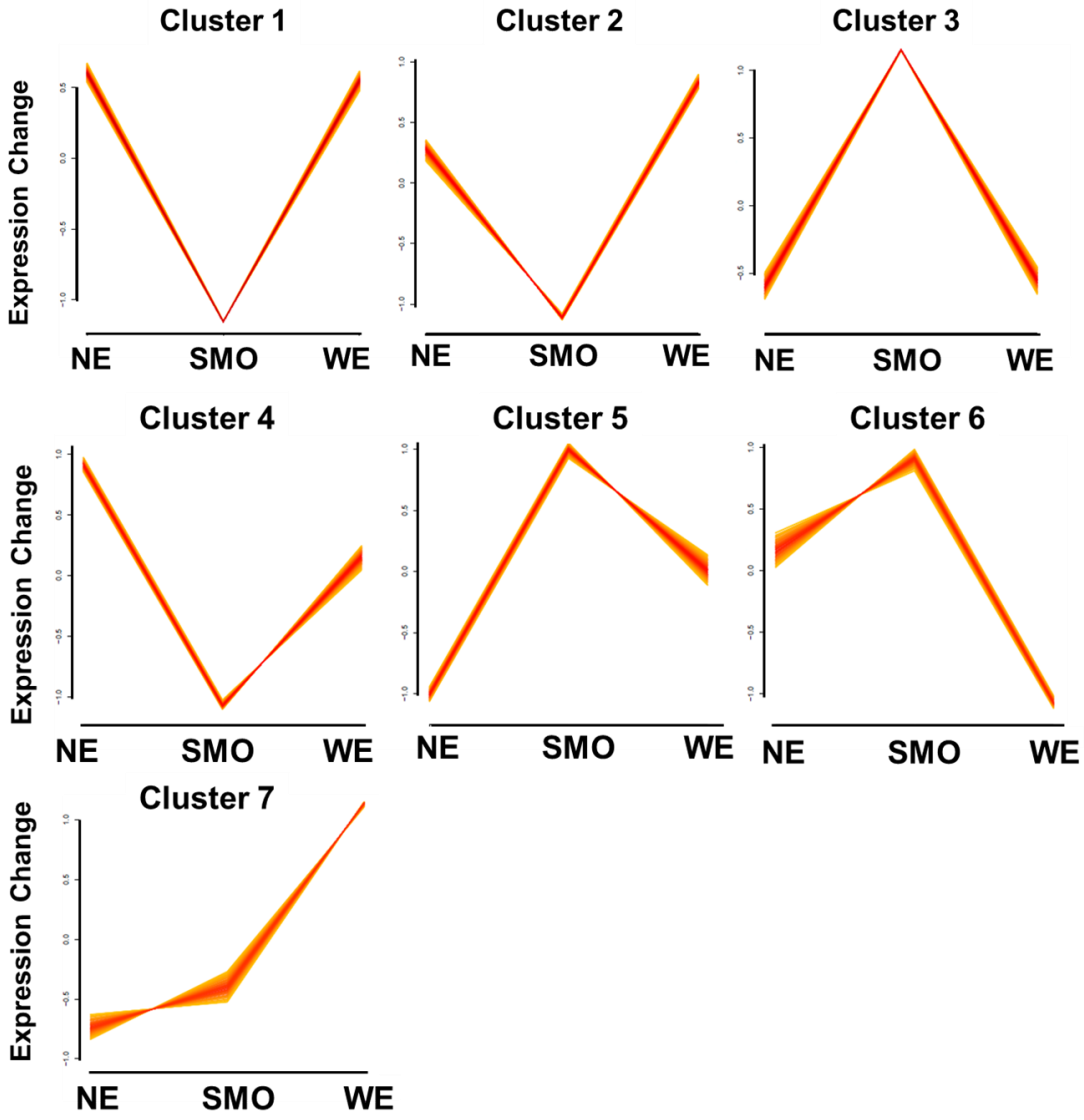

Figure 4.6. Fuzzy c-means clustering to reveal proteomic abundance patterns within the two spatial tissues and the whole embryo.

Spatially distinct energy metabolism within the embryo. While the whole embryo did not show preferential enrichment of any particular metabolic pathways, we found the NE to be enriched in enzymes associated with the TCA cycle and β fatty acid oxidation (compared to the whole embryo and the SMO) whereas glycolysis, pyruvate metabolism, methylmalonyl pathway, and fructose galactose metabolism were enriched in the SMO (compared to the whole embryo and the NE). **Figure 4.7** shows the enzymes and pathways enriched in each of these tissues.

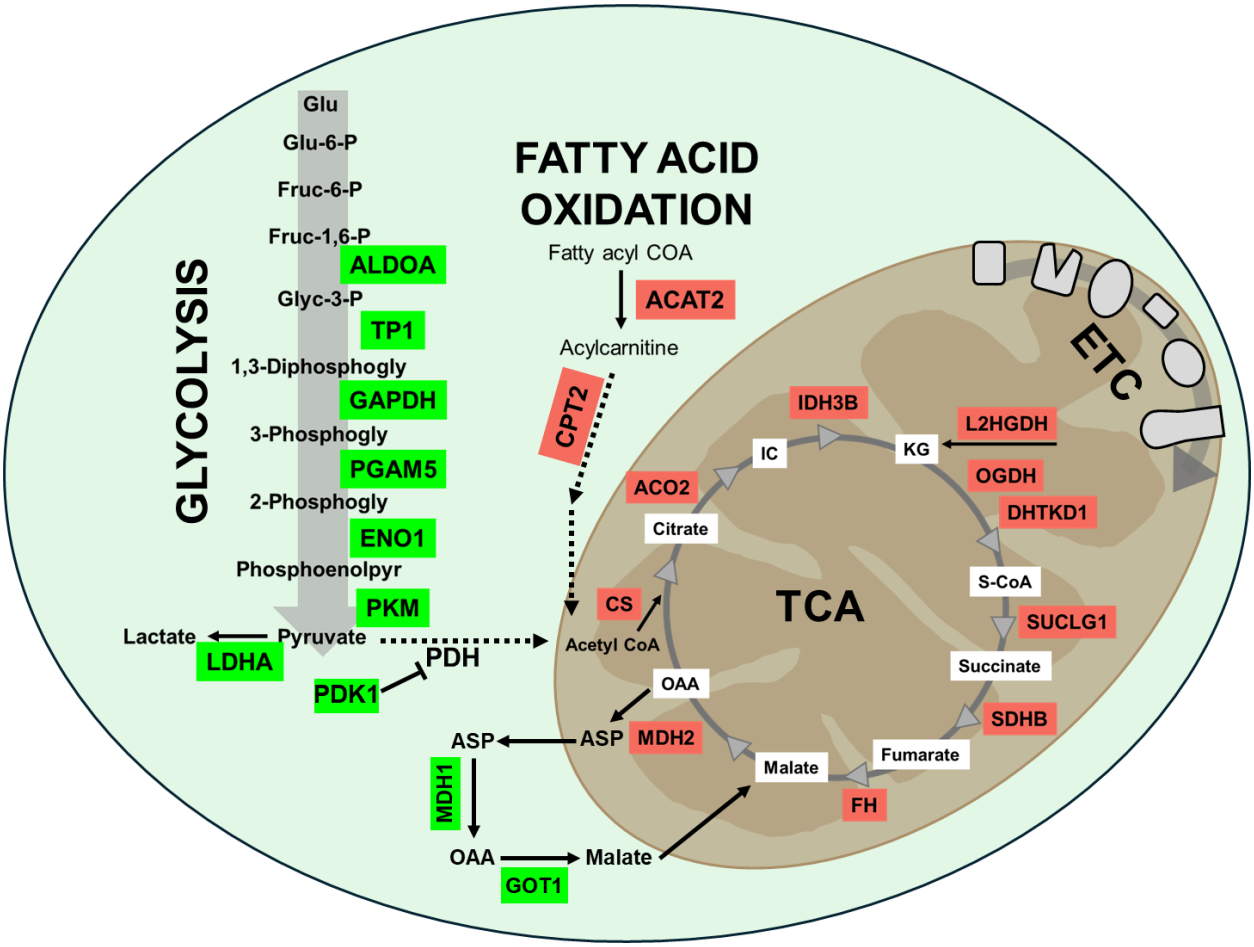

Figure 4.7. Schematic of differentially enriched metabolic pathways. Enzymes involved in the glycolysis pathway were enriched in the SMO while TCA cycle enzymes were enriched in the NE tissue. Full forms of metabolites and metabolic enzymes are listed in the glossary. Color key: green, significantly enriched in the SMO; red, significantly enriched in the NE (compared to the other two tissues)

Most of the enzymes involved in the glycolysis pathway were upregulated in the SMO tissue. The end product of glycolysis is pyruvate which is converted to acetyl coenzymeA via the pyruvate dehydrogenase complex (Pdh) concurrently, the inhibitor of the PDH complex, 3-phosphoinositide-dependent protein kinase-1 (Pdk1) was found to be upregulated in the SMO tissue. Furthermore, we observe an upregulation in the cytosolic enzymes (Got1 and Mdh1) involved in the malate aspartate shuttle which is critical for translocating the electrons to the electron transport chain (ETC) to generate adenosine triphosphate (ATP) as glycolysis produces energy via nicotinamide adenine dinucleotide (NAD and NADH) which cannot cross the inner mitochondrial membrane.[236] These findings indicate glycolysis is active in the SMO and being used as the major energy producing pathway. Surprisingly, we also found lactate dehydrogenase (Ldha) to be upregulated in the SMO cell population, which converts pyruvate to lactate to produce energy through anaerobic glycolysis to compensate for reduced oxidative phosphorylation.

Similar findings were also reported in various models for embryonic patterning. As far back as the 1930s, research employing non-MS methodologies revealed elevated anaerobic glycolysis in the dorsal blastopore lip (the organizer region) in the amphibian embryo.[237] A recent MS based metabolomic study in the *Xenopus* embryos reported similar findings, where the metabolome of the dorsal marginal zone (DMZ, which consists of the organizer) was juxtaposed with the ventral marginal zone (VMZ). The DMZ exhibited high amounts of glycolysis- and glutathione metabolism-related metabolites while the VMZ showcased enrichment of purine metabolism-related metabolites.[238] Similarly, in an investigation involving hESCs, it was found that the addition of exogenous pyruvate could induce differentiation towards the mesodermal lineage by regulating the AMPK/mTOR pathway.[235]

In contrast, the NE tissue showed upregulated expression of the TCA cycle and β fatty acid oxidation related enzymes. As discussed previously, β fatty acid oxidation is regarded as an important pathway for neural differentiation and its dysregulation is known to cause neural birth defects. As the end product of β fatty acid oxidation funnels into the TCA cycle via acetyl CoA, we posit that the NE tissue selectively shifts metabolic programming from glycolysis to oxidative phosphorylation. Such remodeling of the metabolism during neural differentiation has been studied in several stem cell models.[239] Furthermore, a study in neural stem cells reported that a protein called Tigar (TP53 induced glycolysis and apoptosis regulator) fosters differentiation of neural cells through acetyl CoA mediated histone acetylation which in turn controls the translation of proteins involved in neural differentiation.[240] Our proteomic data aligned with the existing literature, as we found significantly higher concentration of Tigar protein in the NE tissue compared to both the SMO and the WE. These encouraging results show a fine balance between energy/metabolic cycles to control the differentiation of embryonic cells during gastrulation. As the metabolic pathways generate energy via ETC, decided to delve deeper into variances in the translation of ETC related proteins to understand mechanisms involved in embryonic patterning.

Role of Energy metabolism via ETC during gastrulation. Based on our proteomic dataset, we conjectured that the NE tissue would demonstrate elevated oxidative phosphorylation (OXPHOS) compared to both the SMO and the WE. In accordance with established literature, we checked the expression of leucine-rich PPR-motif-containing protein (Lrpprc), known to play an important role in ETC complex biogenesis.[241] We found Lrpprc significantly downregulated in the SMO tissue, aligning with our initial hypothesis. Consequently, we directed our focus toward mapping the abundances of mitochondrial proteins associated with the electron transport

chain (ETC). We categorized the protein abundances based on each complex of the ETC. Contrary to our expectations, we did not find upregulation of all complexes within the NE tissue. Although overall both NE and SMO showed upregulation of ETC complexes compared to the whole embryo, we found preferential enrichments of certain complexes within the NE and the SMO. **Figure 4.8** shows the schematic of the ETC complexes and their distribution within the neighboring tissues. The mitochondrial ETC chain consists of 5 main complexes numbered I–V. We analyzed the abundances of various proteins within each complex to evaluate their preferential enrichment within the embryo. The proteomic mapping unveiled that complex II and complex IV were enriched in the NE tissue while complex III and complex V were enriched in the SMO compared to the other two counterparts whereas various subunits of complex I showed mixed upregulation in all three tissue types. Our investigation on individual dissections corroborated our earlier study on pooled tissues within the SMO, which similarly demonstrated enrichment of ETC complex V.

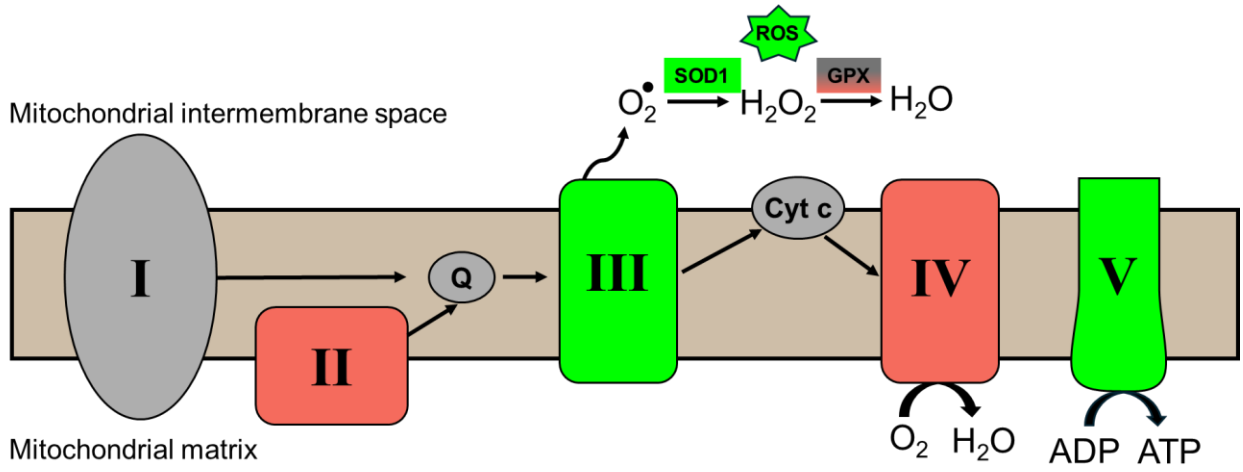

Figure 4.8. Enrichment of ETC complex I–V in the tissues under study. Color key: enriched in SMO, green; enriched in NE, red; no preferential enrichment between the SMO, the NE, and the WE or not detected, light gray; enriched in the WE, dark gray. Abbreviations mentioned in text or glossary.

Another function of the ETC apart from oxidative phosphorylation and ATP generation is the generation of ROS species which participate in various signaling processes. Mitochondrial complexes III–V play an important role in ROS generation. As electrons pass through the ETC complexes, superoxide radicals are generated at complex III which are then neutralized to water at complex IV. Complex V on the other hand acts as an ion channel to pump protons back into the mitochondrial matrix to generate ATP. However, in certain cases, it can reverse the direction and hydrolyze ATP to actively pump protons out of the matrix to reestablish the electrochemical gradient. Complex III is the vital site for the generation of superoxide ions especially in the intermembrane space, ROS radicals escape from the intermembrane space to the cytosol through ion channels.[242] In the cytosol, the radicals are converted to other ROS products such as hydrogen peroxide (H_2O_2) via enzymes such as superoxide dismutase (Sod). H_2O_2 is then neutralized to water through glutathione peroxidases (Gpx).[243] H_2O_2 is a known signaling molecule that has been studied in the context of embryonic development[244] as well as cancer progression[245].

Notably, a recent study also examined the impact of electron transport chain (ETC) complex inhibitors on gastrulation in *Xenopus tropicalis*. The findings indicated a loss of organizer signal at NF Stage 10 in embryos treated with antimycin (a complex III inhibitor).[246] In conjunction, translation of cytosolic Sod1 protein which catalyzes the conversion of superoxide radicals to H_2O_2 was enriched in the SMO compared to the NE and the WE while Gpx related isoforms were depleted in the SMO in our dataset, suggesting spatial enrichment of H_2O_2 within the gastrulating embryo (**Fig. 4.9**). Based on our proteomic results, we hypothesized that enrichment of complex III, complex V, and Sod1 protein in the SMO could indicate heightened ROS generation rather than OXPHOS activity while maintaining OXPHOS in the NE tissue.

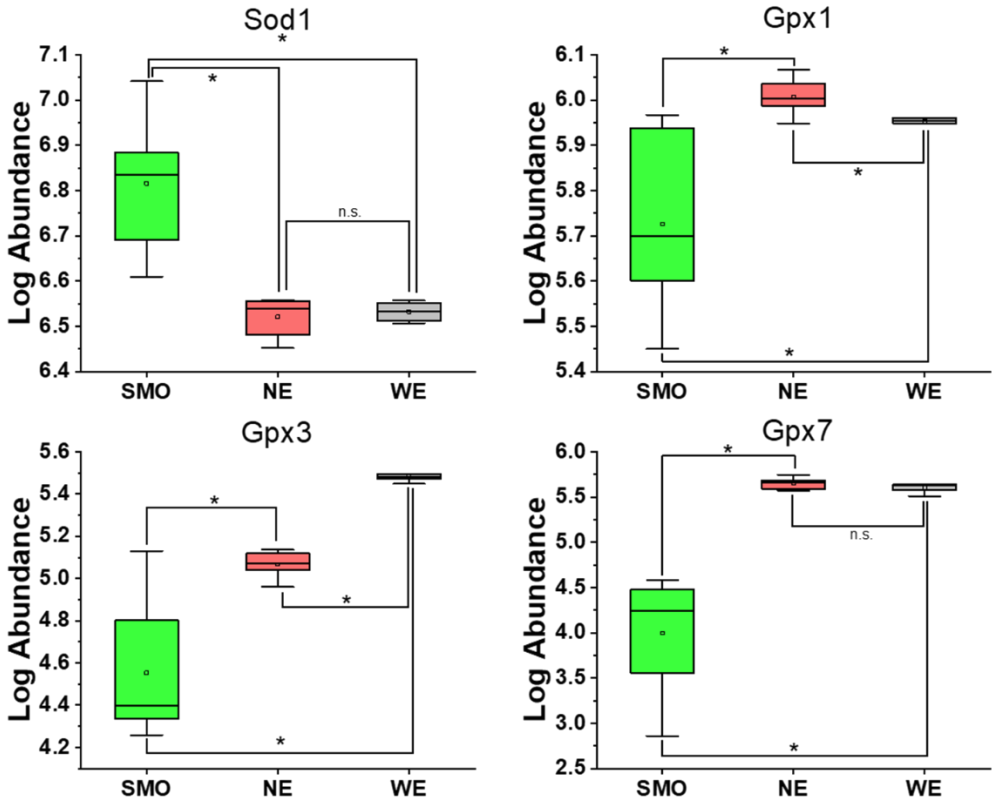

Figure 4.9. Comparison of protein concentation profiles of commonly known enzymes that catalyze ROS metabolism. Abbreviations in text. Color key: enriched in SMO, green; enriched in NE, red; enriched in WE, gray.

Orthogonal validation of proteomic results. Our proteomics results revealed interesting spatial heterogeneity of metabolic pathways. To validate our proteomic results, we opted to employ orthogonal assays to detect the spatial enrichment of small molecules within the gastrulating embryo. To test if ROS is localized spatially within the embryo, we chose to implement fluorescence live imaging. We hypothesize that ROS species get converted to H_2O_2 and localized within the SMO. We used an H_2O_2 live imaging probe to fluorescently label the H_2O_2 within the live embryo. The dye permeates the live cells where it is hydrolyzed by intracellular enzymes to form an impermeable product, this product reacts with H_2O_2 and produces fluorescence (Absorbance λmax= 492 nm, Emission λmax= 516 nm).

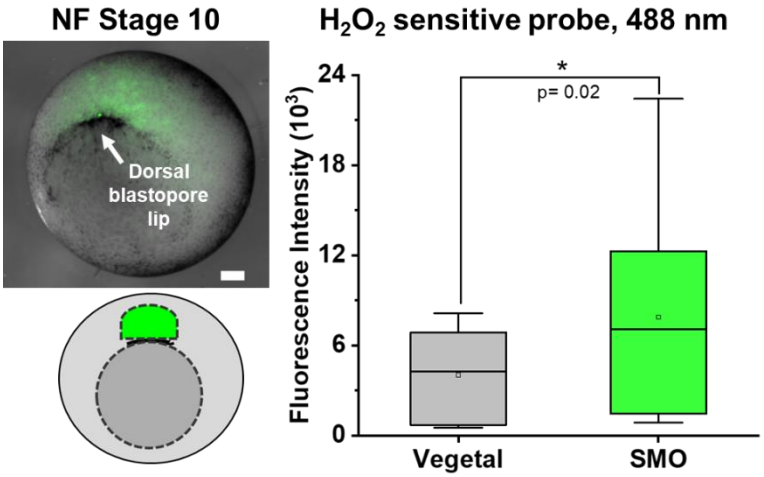

Figure 4.10. Visualization of H_2O_2 localization using H_2O_2 sensitive fluorescent probe in live Stage 10 embryo.The H_2O_2 signal was found to be significantly localized in the SMO region (p=0.02, N=19, scale bar: 100 µm)We incubated the live embryos in the culture media containing a 5 µM probe (see *Methods*) and washed the embryo by nutating in the culture media before imaging. We imaged the live embryos at NF Stage 10 and measured the mean fluorescence intensity within the SMO area and area below the blastopore lip. As shown in **Fig 4.10**, we found statistically significant localized H_2O_2 fluorescence within the SMO region at the beginning of gastrulation, which is marked by the presence of the blastopore lip. According to existing literature, the organizer emerges at the onset of gastrulation and subsequently undergoes involution, culminating in its differentiation into the notochord (mesodermal lineage) by the conclusion of gastrulation.[1] To test if the localization of the H_2O_2 signal coincides with the organizer activity, we traced the H_2O_2 fluorescence intensity temporally during embryonic development. **Figure 4.11** displays the *in vivo* H_2O_2 labeling across multiple stages, including pre-gastrulation (NF stages 8 and 9), gastrulation (NF stages 9.5 and 10), and post-gastrulation (NF stages 11 and 12). By the end of gastrulation, the blastopore lip closes and the mesoderm involutes inside completely, marking the beginning of neurulation.

H_2O_2 fluorescence within the SMO precursor cells was absent in blastula stages prior to gastrulation indicating emergence of H_2O_2 enrichment begins as the organizer is established at the blastopore lip marking the beginning of gastrulation (**Fig. 4.11**). H_2O_2 continues to be spatially enriched within the SMO cell population as gastrulation progresses. At NF Stage 12, weak fluorescence is seen around the blastopore lip, beyond this stage the mesoderm completely involutes and the blastopore lip is closed. Using a fluorescence based assay, we effectively corroborated our proteomic results proving the localization of H_2O_2 within the SMO implying uncoupling of the mitochondrial OXPHOS to generate higher ROS. Moreover, this finding is consistent with the enrichment of glycolysis and increased Ldha (anaerobic glycolysis) seen in the proteomic dataset, which would be essential for energy production in a state of reduced OXPHOS.

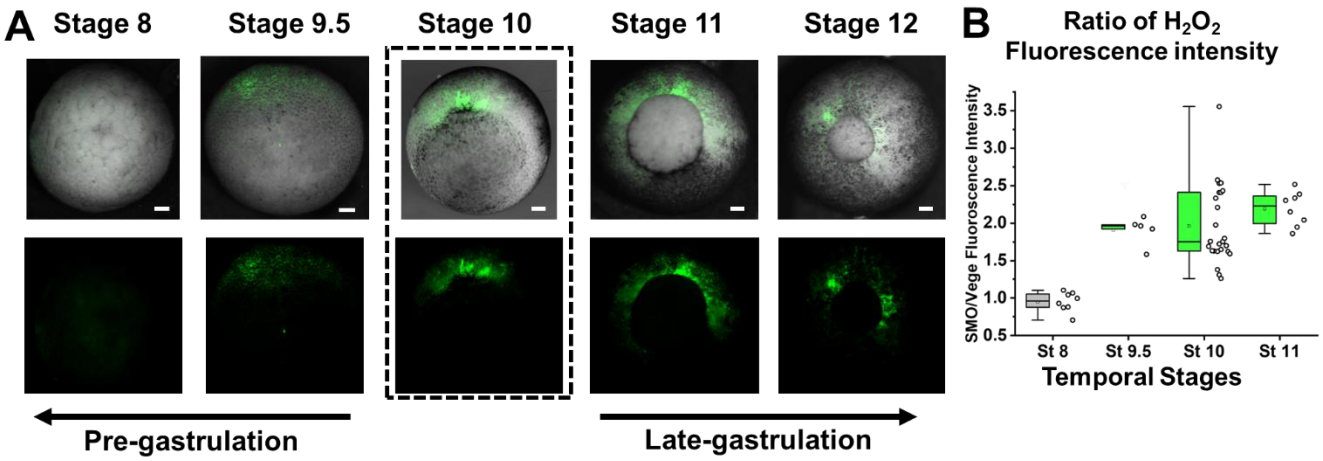

Figure 4.11. Temporal tracing of H_2O_2 signal in the developing embryo. **(A)** H_2O_2 appears at the beginning of gastrulation as the SMO is established and diffuses towards the end of gastrulation. **(B)** The ratio of SMO area to vegetal area (area under the blastopore lip) fluorescence intensity. (Scale bar:100 µm)

Putative role of H_2O_2 in embryonic patterning. Apart from contributing to the mesodermal lineage, the SMO tissue plays an important role in guiding the differentiation of neighboring

cells. Recognized as a dorsalizing center, it controls the induction of neural tissue and dorsoanterior structures within the embryo. To assess if the localized H_2O_2 within the SMO tissue is pertinent for its signaling and patterning activity, we devised morphological assays. Figure **4.12A** demonstrates the experimental strategy to perturb the H_2O_2 enrichment within the embryo. At first, we targeted NF stage 10, as that is when we observed a spike in H_2O_2 accumulation within the SMO region. We perturbed the H_2O_2 gradient by exposing the embryos to H_2O_2 spiked media for a short time. Embryos at NF Stage 10 were incubated in 20% SS spiked different concentrations of H_2O_2 for 2-15 minutes. At exposure beyond 30 minutes, the embryonic development was arrested. The embryos were transferred to normal (unspiked) 20% SS and allowed to develop to larval stages (NF Stage 32/34). The larvae were fixed using the PFA solution (see *Methods*) and imaged under the brightfield microscope. Alternatively, a small clutch of the exposed embryos was immediately used for H_2O_2 fluorescence labeling.

As shown in Figure **4.12 B and C,** 50 mM H_2O_2 and 100 mM (10–15 min) exposure disturbed the H_2O_2 gradient as depicted by a significant decrease in the ratio of fluorescence intensity of the SMO area to the area under the blastopore lip (Mann Whitney test, $p < 0.05$, N=12–25). As the organizer is known to guide the dorsoanterior and neural induction, we evaluated the three parameters comparing control larvae with those exposed to H_2O_2: length of the spine, diameter of the eye, and area of the cement gland. The eye and spine were chosen as they are derived from the neural cells, with the eye additionally situated in the anterior region. The cement gland was selected as it marks the anterior most region in the larvae which develops from ectodermal cells and lies at the border of neural (dorsal) and epidermal (ventral) fates.

At shorter exposure times, the larvae from the experimental group (50 mM H_2O_2, 2 min exposure) were indistinguishable from the control while gradual perturbation was seen in larvae

exposed for 8 min. The larvae exposed to 50 mM H_2O_2 media for 15 min or 100 mM H_2O_2 media for 10 min showed significant defects in embryonic patterning (**Fig**. **4.12D**). The spine lengths, eye diameter, and cement gland area were significantly smaller in the experimental group compared to the control group (Mann Whitney test, $p<0.05$, N=15–35) indicating dysregulated neural and dorsoanterior development (**Fig**. **4.12E**). The embryos exposed to H_2O_2 exhibited dysregulation akin to the ventralization induced by ultraviolet (UV) radiation, a well-established method for inhibiting organizer activity in embryos.[247] Although UV-irradiated embryos undergo gastrulation and develop the three germ layers, they fail to form dorsal structures instead, only the ventral mesoderm (devoid of the notochord) and epidermis are produced. Severely ventralized embryos resemble belly pieces.[1] The phenotypic similarity observed between H_2O_2 perturbation and UV ventralization further confirms that the spatial H_2O_2 gradient within the embryo plays a significant role in embryonic patterning.

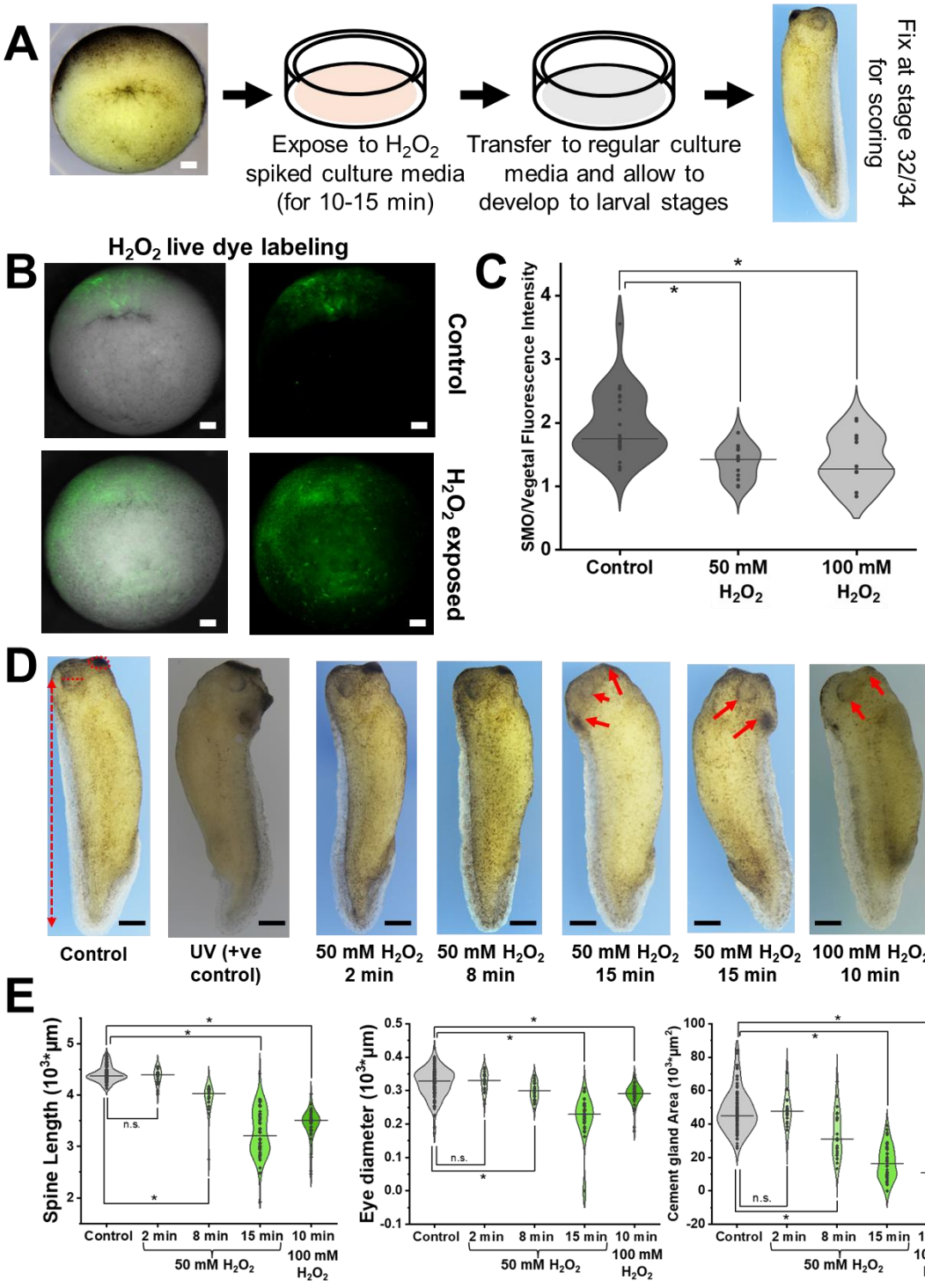

Figure 4.12. The phenotypic effect of perturbing H_2O_2 gradient. **(A)** Experimental strategy to expose embryos to H_2O_2 spiked media to dysregulate H_2O_2 gradient. **(B)** Validation of disturbed gradient using H_2O_2 fluorescence assay **(C)** Statistically significant suppression of H_2O_2 enrichment in the SMO in experimental group compared to controls evaluated by measuring the ratio of fluorescence intensity between the SMO area and the vegetal area. **(D)** Perturbation of dorsoanterior and neural development compared to controls with increasing exposure to H_2O_2.

(E) Measurement of spine length, eye diameter, and cement gland area of control and experimental groups indicating statistically significant defects in spine, eye, and cement gland (Mann Whitney test, $p < 0.05$, n.s.= non-significant, N=15–35; scale bar: 100 µm (white), 500 µm (black)).

We further examined if the effect of H_2O_2 perturbation was specific to the temporal evolution of the organizer and neural induction. We speculated that if the H_2O_2 gradient was pertinent for the organizer activity, perturbing it before the organizer is established at blastula stages (NF Stage 8/9) would dysregulate the patterning similar to the exposure at NF Stage 10, however, disturbing the gradient after completion of gastrulation when the SMO has already established the neural ectoderm and undergone involution to form the notochord (mesoderm lineage) should have little to no effect on the patterning. To investigate the temporal specific effect of H_2O_2 exposure, embryos were exposed to H_2O_2 spiked media at Stage 8 (before SMO is established), at Stage 13 (end of gastrulation, beginning of neurulation), and Stage 20 (towards the end of neurulation). Embryos were exposed for 10–15 min and then cultured in regular unspiked media and allowed to develop into larvae for scoring. **Figure 4.13** displays the results from the temporal dysregulation of the H_2O_2 gradient. As anticipated, embryos exposed to H_2O_2 before gastrulation showed significant aberrant phenotypic features (spine, eye, and cement gland) which could be attributed to the disturbance in the organizer establishment and function. Conversely, embryos exposed post gastrulation were unaffected by H_2O_2 exposure. Embryos exposed at Stage 13 exhibited slight dysregulation (only spine lengths shorter) whereas embryos exposed at Stage 20 resembled controls showing no signs of defective patterning. This pilot study opens exciting avenues to study how H_2O_2 controls signaling within the embryo thereby expanding our existing knowledge on the organizer function and mechanisms involved in tissue induction and differentiation in normal embryos.

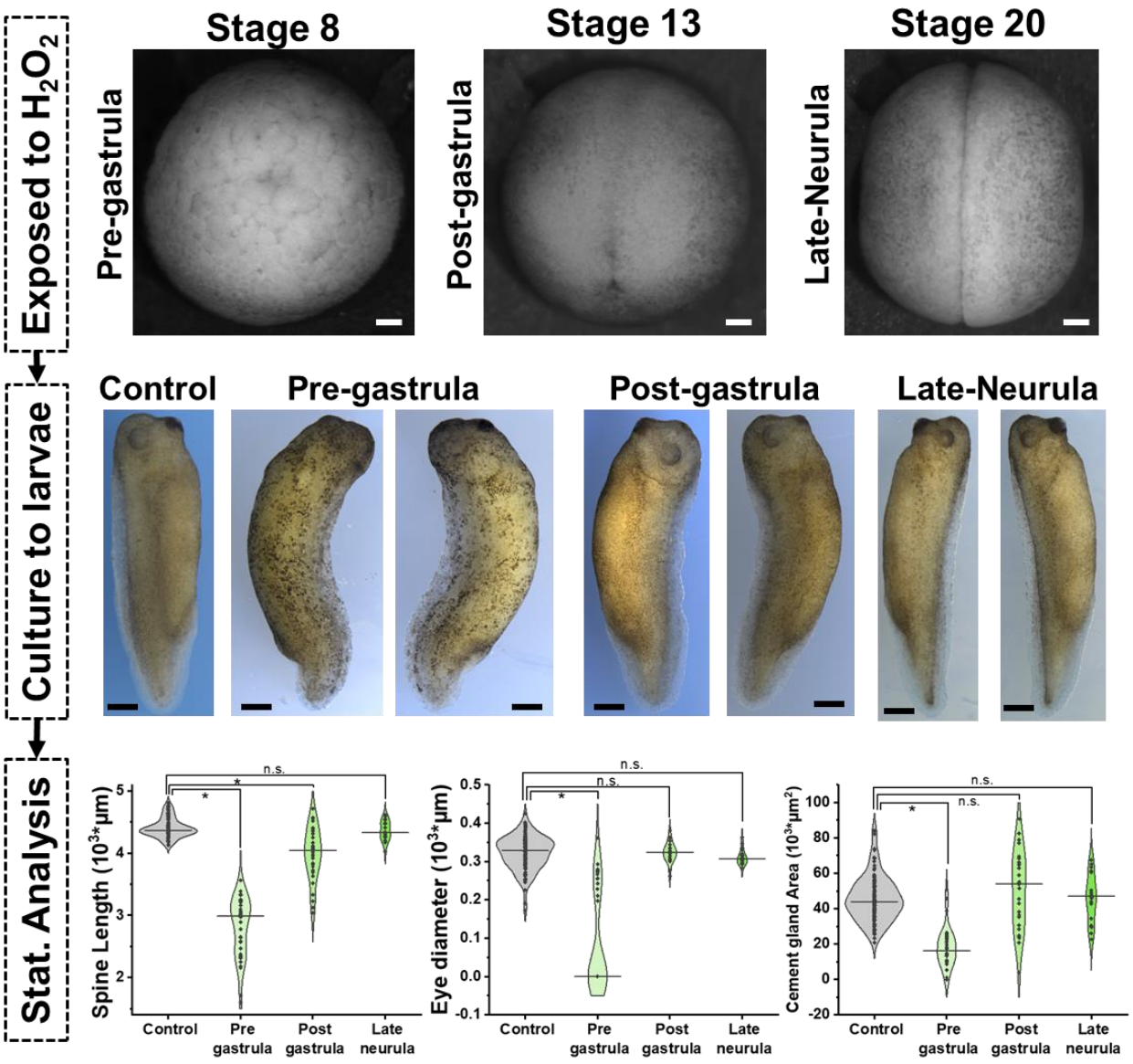

Figure 4.13. Phenotypic evaluation of temporal dysregulation of H_2O_2. The top panel shows stages of embryos at the time of H_2O_2 exposure (100 mM, 10 min). Observed phenotypic changes in larvae (control, unexposed; experimental, exposed at different developmental stages) are displayed in the middle panel. Comparison of dosroanterior and neural features between control and experimental groups (bottom panel) using Mann Whitney non parametric test ($p < 0.05$, n.s.= non-significant, N=15–35; scale bar: 100 µm (white), 500 µm (black)).

Preliminary findings from spatial metabolomic analysis. To orthogonally corroborate the differences in glycolysis and OXPHOS between the SMO and the NE tissue, we aimed to analyze metabolites in these tissues such as the glycolysis end products (pyruvate and lactate) and TCA cycle metabolites. Small organic acids are often difficult to retain on LC columns, hence we opted to perform metabolic analysis using GC-MS. We tailored the previously established method[248] by our collaborator for our cell lineage dissections of the SMO and the NE. We pooled 20 dissections of the SMO and the NE tissue and 2 WE for metabolomic analysis. The samples were lysed to extract the metabolites in an organic solvent, isotopic internal standards of each metabolite under study were spiked in a known amount to the extract and the mixture was derivatized for GC-MS as described elsewhere.[249] A targeted MS method was used to selectively analyze pyruvate, lactate, and TCA metabolites.

Due to sensitivity limitations, the detection of all targeted metabolites could not be accomplished from pooled tissues. Nonetheless, pyruvate was identified in all three tissue types, with its average abundance found to be higher in the SMO compared to the NE and the WE which aligns with our proteomic findings. These initial findings offer promising results to refine our analytical method for metabolomic analysis further. Future experiments will prioritize improving the detection sensitivity of organic acids and include all glycolysis metabolites to achieve a complete glycolysis and OXPHOS profile of these tissues under study. Furthermore, metabolic flux analysis using isotopic standards can further reveal the difference in the rate of glycolytic and TCA metabolic pathways. In conjunction with MS based assays, we anticipate tailoring colorimetric or luminescence assays which are typically used to measure glycolysis and OXPHOS in cell suspensions for our embryonic tissues.

4.5 Conclusions.

In this study, we demonstrated the application of our ultrasensitive MS proteomic approach to discover proteomic gradients within the vertebrate embryo during gastrulation. The overall goal of this study was to generate a deep proteomic map to uncover spatial proteomic heterogeneity within two neighboring regions known to give rise to different tissue lineages post gastrulation, namely the neural and mesodermal tissues. Additionally, the mesodermal fated cell population (also known as the SMO) acts as an important transient signaling center at the beginning of gastrulation coordinating neural induction and guiding embryonic patterning. Thus, proteomic profiling of the SMO along with the induced NE tissue is of key research interest to understand healthy embryonic development.

Our HRMS proteomic approach detected over 8000 proteins in these spatially distinct embryonic regions and ~3500 proteins showed significantly different expression patterns among the three tissue types (the SMO, the NE, and the WE). Enrichment analysis revealed differential regulation of several canonical pathways corroborating our proteomic dataset with published transcriptomic and proteomic studies. The dataset also revealed differences in metabolic pathways such as glycolysis and TCA cycle within the SMO versus the NE tissues. Our proteomic dataset further led to the discovery of preferential enrichment of certain ETC complexes within the SMO and the NE, highlighting the role of ROS and OXPHOS in tissue induction and dosroanterior patterning. Using molecular fluorescent probes we orthogonally confirmed spatial enrichment of ROS byproduct, H_2O_2 in the organizer.

We developed a functional assay in the developing embryo to assess the importance of H_2O_2 localization in the SMO. Spatiotemporal dysregulation of H_2O_2 gradient led to developmental

defects revealing a novel role of H_2O_2 as a signaling molecule to guide neural tissue induction and dorsoanterior patterning in the vertebrate embryo. Moving forward, future research can build upon this novel finding to elucidate the precise mechanisms through which ROS influence SMO functions and NE induction, ultimately advancing our understanding of vertebrate embryogenesis and potentially offering new targets for therapeutic interventions or regenerative medicine approaches. Subsequent studies will focus on finding the downstream protein targets of H_2O_2 and continue metabolic flux analysis to delve into the significance of the varied enrichment of glycolysis and the TCA cycle within these particular tissue lineages. Furthermore, a comprehensive spatiotemporal proteomic mapping spanning from early embryonic to the conclusion of gastrulation stages will aid in delineating the process by which embryonic cells become committed to their respective tissue lineage.

Acknowledgements. Parts of this work were funded by the National Institute of General Medical Sciences (award 1R35GM124755 to P.N.), the University of Maryland-National Cancer Institute Partnership Program (seed award to P.N.), and the COSMOS Club Foundation (fellowship award to L.R.P.). We thank Annie Liu for assisting with imaging experiments and metabolomics sample preparation. We thank Bowen Shen for conducting UV ventralization experiments. We express our gratitude to Dr. Nishanth Sunny for sharing his expertise in GC-MS metabolomics and permitting the use of his laboratory facility and instruments for GC-MS analysis.

Chapter 5: Spatiotemporal Proteomics of the Branchial Arches

This chapter is based on the Pade, L.R., Yoon, J., Daar, I.O., Nemes, P. (2024) manuscript draft in preparation.

Contribution: L.R.P. prepared the samples for proteomic analysis, analyzed the samples via LC-MS, performed data analysis, interpreted the proteomic results, drafted the manuscript, and revised the contents.

5.1 Abstract

During vertebrate embryonic development, embryonic cells compartmentalize spatially to induce tissue formation. In vertebrates, a group of stem-like cells called neural crest cells (NCCs) dynamically interpret and respond to diverse signals to migrate significant distances and differentiate into many tissues. A subset of NCCs called cranial neural crest cells (CNCs) migrate spatially to give rise to four branchial arches (BAs). Each BA forms different structures in the head and jaw region during vertebrate embryonic development. How the transcriptomic and proteomic expression patterns differ between these four BAs to form different tissues remains poorly understood. Knowledge of proteins expressed in these cells will improve our understanding of how stem cells divide and differentiate into different craniofacial structures. Using our previously developed YODEC approach we enabled the detection of ~9000 proteins from the four CNC streams. The new approach doubled the coverage compared to traditional multiplexing precluding the need to pool large sample amounts for yolk depletion and minimizing sample loss during processing steps. We detected several important proteins such as TBC1d24, Ror2, cadherins, and ephrin-type receptors known to play a role in the migration and differentiation of CNCs.

5.2 Introduction

Cranial neural crest cells (CNCs) are a subset of multipotent progenitor cells derived from neural crest cells unique to vertebrates which give rise to various tissues during development. CNCs specifically give rise to craniofacial structures such as the jaw, skull, cartilage, melanocytes, and ganglia of the peripheral nervous system. The neural crest originates from the boundary area of the neural plate, situated between the neural plate and the neighboring non-neural ectoderm. Within this border region, neural crest cells are determined through a combination of inducing signals that commence during gastrulation. Throughout neurulation, they remain positioned at the border of the neural plate as it evolves into the emerging neural folds. Subsequently, following the closure of the neural tube, they migrate to a region within the dorsal neural tube. The precursor population of the neural crest expresses a distinctive set of transcription factors, which include Snail2 (Slug), Sox10, FoxD3, and Sox9, collectively referred to as neural crest specifier genes.[250] These cells undergo three phases namely induction, migration, and differentiation. The nonneural ectodermal cells are first induced and then undergo epithelial to mesenchymal transition (EMT) to adopt migratory behavior as they detach from other neuroepithelial cells, and eventually differentiate to give rise to different tissues. Due to their unique migratory behavior and ability to induce different tissue types, these cells are often referred to as the fourth germ layer and are used as a key model for studying cell differentiation as well as cancer metastasis and progression.[251]

In the *Xenopus* model, the CNCs are induced during early gastrulation and undergo EMT once the neural tube closes to form the four streams of cells called the branchial arches (BAs), which are labeled BA1–BA4. The first arch is known to form parts of jaw cartilage (Meckel and palatoquadrate) and subocular cartilage. The second arch forms the cartilage around the hyoid bone and the third and the fourth arch form parts of the gills.[252] Initially, the CNCs migrate from

the dorsal to ventral (DV) axis, followed by migration along the anterior to posterior (AP) axis as they differentiate into four streams. The migration along the DV axis has been extensively studied in various models[253-255], however, only a few studies have focused on understanding the migration along the AP axis. Understanding how the transcriptomic and genomic expression patterns change along the AP axis across the four BAs can help identify unique pathways or markers critical for BA differentiation and induction of different craniofacial tissues.

A previous study from the Daar research group explored the role of Ephrin signaling along the AP axis as CNCs migrate. Ephrin B2 was found to interact with TBCd124 to guide the migration of the CNC streams. It was observed that Ephrin B2 interacts with TBCd124 to facilitate the guidance of CNC stream migration. Notably, when TBCd124 was knocked down, the migration of BA2–BA4 was disrupted, while the migration of BA1 remained unaffected, thus raising the possibility of distinct signaling patterns along the AP axis.[256] Apart from transcriptomics, complementary knowledge of proteomic expression patterns within these tissues can provide a holistic understanding of mechanisms regulating craniofacial development. At present, there are no proteomic studies published in the literature specifically targeting the four branchial arches.

To bridge this knowledge gap, this study aimed to undertake an extensive spatial proteomic analysis of the four branchial arches. Along with spatial differences, we also sought to compare the temporal differences between pre-migratory CNCs and post-migratory CNCs (or BAs). Deep proteomic investigations typically necessitate several hundred micrograms of initial protein quantities. However, each branchial arch comprises approximately 2,000 cells, corresponding to around 200 ng of proteins, with only approximately 20 ng representing non-yolk proteins in the *X. laevis* embryo. In pursuit of this objective, we sought to utilize analytical methodologies previously established to augment the depth of coverage within these cellular populations.

5.3 Experimental section

Materials and Reagents. LC-MS grade solvents and reagent-grade chemicals were procured from Fisher Scientific (Pittsburgh, PA) or Millipore Sigma (St. Louis, MO) unless explicitly stated otherwise. Reagents for lysis buffer include SDS, combretastatin A-4, cytochalasin D, and nonidet P-40 substitute (NP-40). Proteomics reagents include trypsin protease (Pierce, MS grade), TMT 6plex isobaric label reagent set, DTT, and IAD.

Solutions. Embryological solutions include 1× Modified Barth's Saline (MBS) for culturing and 1× MBS containing 0.1% gentamycin for dissections prepared according to established protocols for embryo culture and tissue dissection.[257] Yolk depletion buffer was prepared as described previously (see Chapter 3) following publications.[159, 183]

Animal Care and Embryology. Embryology and sample collection were carried out at the NCI campus. All animal procedures reported in this study that were performed by NCI-CCR affiliated staff were approved by the NCI Animal Care and Use Committee (ACUC) and per federal regulatory requirements and standards. All components of the intramural NIH ACU program are accredited by AAALAC International. Xla.Tg(sox10: GFP)Jpsj transgenic frog line (males and females) was generated by the National Xenopus Resource (NXR). Embryos were obtained by inducing the females using gonadotropin hormone to lay eggs followed by *in vitro* fertilization using established protocols.[201] After obtaining the embryos, the jelly coat was dissolved using 2% (w/v) cysteine solution (pH 8) and transferred to a sterile 1× MBS solution. Embryos were allowed to develop to NF Stage 17 (pre-migratory CNCs) or Stage 26 (post-migratory CNCs or BAs).

Tissue Collection. Sox10: GFP transgenic embryos allow lineage tracing of CNCs during embryonic development. CNC cells appear as a single cluster in the pre-migratory stages and branch into four streams of BAs (BA 1–4) at the tailbud stage (Stage 26). For the spatial study, each arch was dissected at Stage 26 under the stereomicroscope using fluorescence and frozen in a separate vial. 10 dissections of each arch were pooled to produce 1 BR. A total of 6 BRs were collected for the spatial study. Additionally, a large pool of all BAs was collected to serve as a carrier for the YODEC strategy. For the temporal study, 10 dissections of pre-migratory CNCs (Stage 16) were pooled for one sample and a mix of all four BAs (Stage 26) was used as the post-migratory CNC sample.

Sample Preparation. Individual arches (BA 1–4) were lysed using the conventional lysis buffer (as described in Chapter 2). The pooled BAs (carrier) were yolk depleted as described previously. For temporal analysis, the pre-migratory CNCs and post-migratory CNCs (mixed pool of BAs) were yolk depleted in vials. All samples were reduced using DTT, alkylated with IAD, and digested using trypsin (1:50 enzyme to sample ratio). For spatial analysis, individual BAs and the BA carrier were tagged with isobaric TMT reagents and combined. The ratio of total proteome digest in analytical channels to the YODEC channel was 1. Each TMT set contained one BR. Each BR was fractionated offline into 75 fractions at high pH and concatenated to 8 fractions. The fractions were dried down and reconstituted in aqueous 0.1% FA for LC-MS analysis.

Mass Spectrometry. The µPAC RPLC separation gradient was the same as outlined in Chapter 3. For the spatial study, the TMT SPS MS^3 multiplexed quantification method was employed with the following parameters: MS^1 at 120,000 FWMH resolution in the orbitrap analyzer; scan range, 400–1600 Da; Max IT, 50 ms; AGC target, 4×10^5 counts; microscans, 1; monoisotopic

peak determination, peptide; dynamic exclusion, 60 s. For data dependent acquisition, the precursor ions were isolated with a 0.7–Da window and initially fragmented with CID at 35% NCE (helium) with an intensity threshold of 5×10^3 counts. The MS^2 spectra were acquired in the ion trap (AGC target, 1×10^4 counts; max IT, set to auto mode; microscans, 1). The top 10 MS^2 fragments were isolated using synchronous precursor selection (MS^3) with a scan range of 100–500 Da and fragmented with HCD at 65% NCE (nitrogen). The MS^3 spectra were acquired at 15,000 FWMH in the orbitrap (AGC target, 5×10^4 counts; max IT, auto mode; microscans, 1). Samples under temporal study were analyzed using label-free quantification using the MS method described in Chapter 3.

Data Analysis. Raw HRMS were processed in Proteome Discoverer (version 2.2, Thermo Scientific) with SEQUEST searching against the *Xenopus laevis* proteome downloaded from UniProt containing 60,000 entries. Search parameters included trypsin as the enzyme, allowed up to 2 missed cleavages, cysteine carbamidomethylation, and TMT label as fixed modification, methionine oxidation as a dynamic modification, all searches included a minimum of 1 unique peptide, 10 ppm mass tolerance for precursor masses, 0.6 Da mass tolerance for MS^2 fragment masses, and 20 ppm mass tolerance for TMT labels. Peptides and proteins were filtered with <1% false discovery rate (FDR). Unique and razor peptides were used for quantification. Proteins were quantified based on their MS^3 reporter ion abundances (for spatial analysis) or LFQ abundances (for untagged temporal samples). Data filtering and normalization were performed in MetaboAnalyst 3.0[203] or Perseus[204]. For the spatial study, TMT batches were normalized using an online platform called BatchServer[258] to reduce batch effects followed by log-transformed for further evaluation. GO was annotated in the PANTHER classification

system (version 16.0). Statistical overrepresentation test and fuzzy mean clustering were performed using PANTHER and VsClust respectively.[187]

5.4 Results and Discussion

Deep spatial proteomics of BAs. The CNC streams develop post gastrulation at later stages of embryonic development spanning from Stage 16–26 (**Fig. 5.1A**). Our first objective was to evaluate yolk interference in these cell populations. Given that yolk serves as a source of energy and is consumed over the course of embryonic development, it was pertinent to determine the yolk content in CNCs to tailor the analytical approach. To this end, we performed standard bottom-up proteomics on pre-migratory CNCs and post-migratory CNCs. In both stages of CNCs, ~70% of total protein abundance belonged to the yolk proteins, with hundreds of yolk PSMs eluting throughout the separation window indicating the need for yolk depletion (**Fig 5.1B**). Next, we evaluated if the CNCs/BAs can be yolk depleted in a vial or using µYODE. Even with a pool of 100 dissections of BAs, the yolk pellet was extremely small to separate the yolk-free supernatant without contamination effectively (**Fig 5.1C**). Therefore, we chose to utilize the YODEC approach for the BAs.

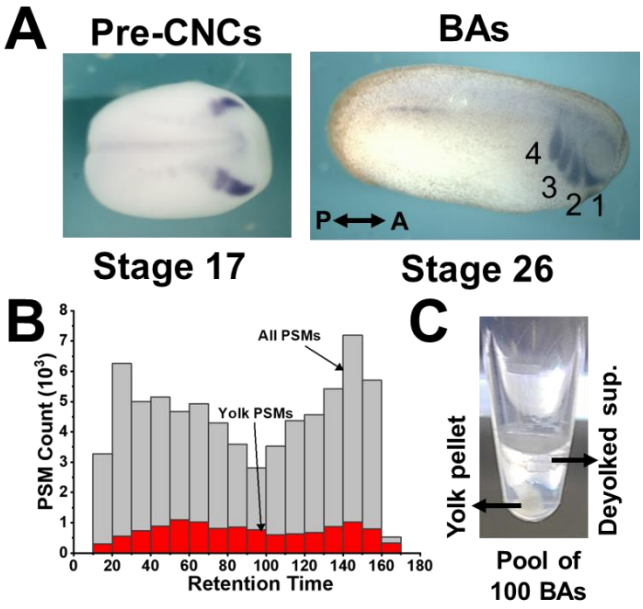

Figure 5.1. Deep proteomics of BAs. **(A)** Tracing temporal evolution of CNCs to BAs using whole-mount in situ hybridization (Ish) with CNC marker Twist. **(B)** Yolk PSMs in BAs binned over the retention time window. **(C)** Yolk depletion of 100 pooled BAs displaying extremely small yolk pellet. Key: A, anterior; P, posterior; sup., supernatant

Figure 5.2 displays the proteomic workflow for spatial BA analysis. Each BA stream was dissected using sharp forceps and frozen in a vial. 10 dissections of each stream were pooled together to prepare 1 BR. A total of 6 BRs of the four BAs were collected. A large pool of all BAs was used for the preparation of the carrier. Individual BA replicates were processed using standard bottom-up proteomic steps without yolk depletion. The large BA pool was yolk depleted to prepare the carrier for the YODEC strategy. All the samples were reduced, alkylated, digested, and barcoded with TMT tags. The individual four BAs were combined in a 1:1 ratio while the YODEC (carrier) was spiked in a higher amount. The resulting tagged mixture was fractionated offline and each fraction was analyzed using the (μPAC) LC-MS^3 method (see *Methods*). Each TMT batch consisted of 1BR and a total of 6 TMT batches were analyzed.

To decide the deyolked carrier amount to spike into the biological samples, we conducted a preliminary analysis with different carrier amounts where the deyolked carrier was mixed in 1×, 2×, and 4× ratio (ratio of carrier peptide digest amount to total peptide digest in analytical channels). Unlike the early-stage embryo, YODEC for BAs showed a steep dilution at 1× and plateaued at higher amounts (**Fig. 5.3**). We chose the 1× ratio for the deyolked BA carrier as it gave similar yolk dilution and lesser missing quantification values in biological channels compared to 4× condition. We identified over 8000 proteins from the 6 biological replicates. We filtered the data to extract the quantitative information, ~5000 proteins had quantitative protein expression abundances in at least 3 of the 5 biological replicates. We further performed normalization across all 5 TMT batches. A large number of TMT batches often introduce batch effects (variances in quantification due to different experimental parameters) which mask the true biological variances. We employed an online proteomics data normalization tool (BatchServer) to evaluate and remove batch effects across different biological replicates.

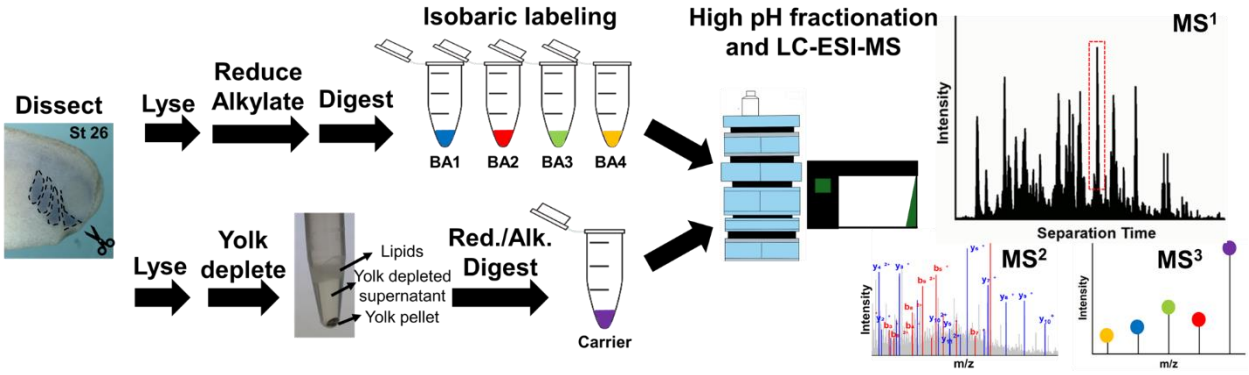

Figure 5.2. Analytical workflow for spatial BA analysis using YODEC. (The Ish image is used for a clear demonstration of the four arches, the sample collection used the Sox10 GFP transgenic line.)

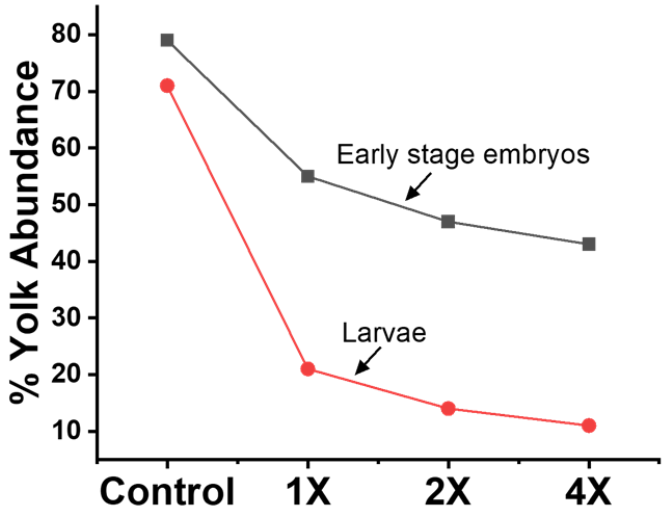

Figure 5.3. YODEC (carrier amount) optimization for BAs collected from Stage 26 (larvae).

Hierarchical clustering to uncover spatial differences. We performed unsupervised hierarchical clustering to reveal proteins that are significantly different between the four BAs. **Figure 5.4** shows a heatmap of the top 75 proteins and the differences in their fold changes within the four BAs. remarkably distinguishing all four BAs. Especially, the proteomic translation in BA3 was significantly different than the other three BAs. We surveyed the literature to find community published corresponding genomic expression data or Ish libraries for the proteins of interest using Xenbase.[12] We found several published Ish images that matched our proteomics data as displayed in **Figure 5.4**. For example, Mthfr (methylenetetrahydrofolate reductase) and C3.L (complement C3) show the lowest protein concentration in BA3 which matches the localized RNA expression visualized using Ish. In CNCs, the significance of 5-methyltetrahydrofolate has emerged, demonstrating its ability to alleviate alcohol-induced CNC migration defects[259] while complement C3a has been highlighted as crucial for coattraction and coordinated movement of CNCs to prevent excessive dispersion of cells after EMT.[260] Similarly, both mRNA and protein expression of Rpe65 (retinal pigment epithelium-specific 65 kDa protein) and Denr (density regulated re-initiation and release factor) are elevated in BA1 and

BA4 and lower in BA2, and lowest in BA3. Rpe65 is reported as an anterior CNC marker[261] while Denr has been studied for its role in cancer progression[262]. Gsn (Gelsolin) is observed to be enriched in BA1 compared to other arches and shows a similar Ish pattern and its expression is recognized in the pharyngeal arches/branchial arches. These examples provide compelling evidence to support our proteomics data. Given the absence of published studies on the critical role of differential expression of these protein markers in BA differentiation and craniofacial development, this discovery paves the way for exciting future research opportunities.

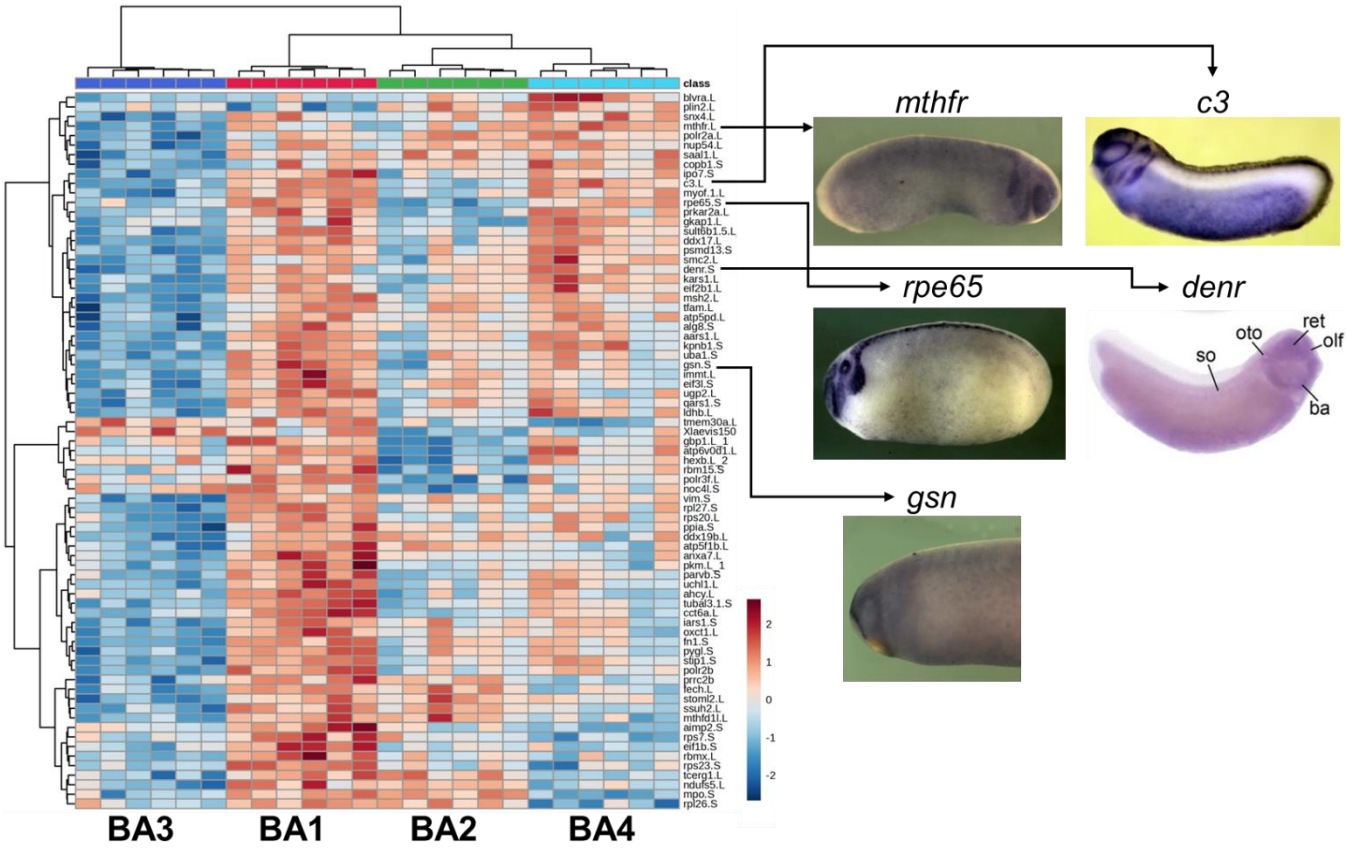

Figure 5.4. Hierarchical clustering distinguishes four BAs based on the top 75 proteins. Examples of select proteins comparing RNA expression (Ish) with protein concentrations showing similar trends across four BAs. Ish images were downloaded from Xenbase and reproduced from the following sources (XDB3 web data catalog of the NIBB/NIG/NBRP *Xenopus laevis* EST project, CNRS UMR 8080, and Grant PA et al. (2014)).[12, 263]

We performed STRING analysis on the top 75 proteins to find correlations or interactions between these proteins. **Figure 5.5** shows the string map of the top 75 proteins from the heatmap that had known interactions published in the literature. The proteins belonged to the following GO categories: translation, transcription regulation, mitochondrial/ETC related proteins, transporters, folate and homocysteine metabolism, protein folding, glycogen metabolism, and EMT markers. The roles of some of these proteins in craniofacial development have been emphasized in the literature. For example, Rbmx is found in neural crest cells and its knockdown using morpholino in *Xenopus* caused eye abnormalities, jaw irregularities, absence of the trigeminal ganglion, and decreased melanocyte count in the trunk during tadpole stages.[264] Similarly, Vim and Fn1 genes are known to play a crucial role in cell migration and EMT and have been detected in several studies.[265] Importins (Ipo7 and Ipo8) were demonstrated to be critical for odontoblastic differentiation which arise from CNCs[266].

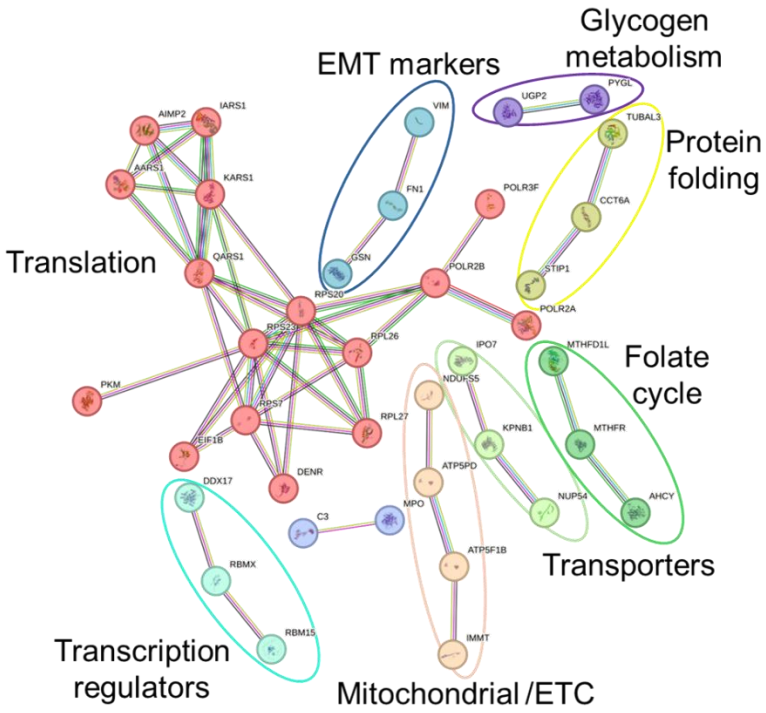

Figure 5.5. STRING analysis of heatmap proteins. Full forms are listed in the glossary.

Spatial proteomic gradients across the four arches. We utilized ANOVA statistical analysis to identify significant differences in protein expression among the four arches. We found ~900 statistically significant proteins that were differentially regulated between BA1–BA4. We performed fuzzy c-means clustering on ANOVA significant to distinguish expression profiles and categorize proteins into various expression patterns observed across the arches. **Figure 5.6** shows the abundance profiles of different proteins grouped into four clusters. Proteins under each cluster were then tested using a statistical enrichment test to reveal significantly enriched signaling pathways to biological processed again *Xenopus* or human database using PantherDB.[267]

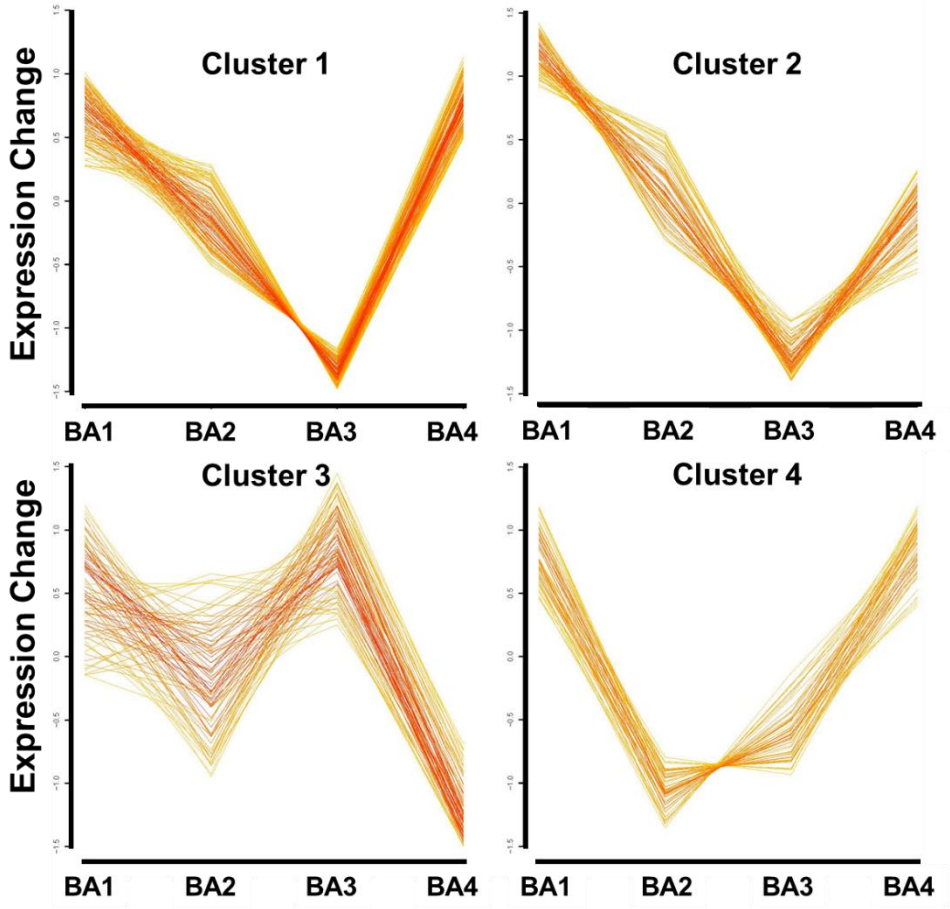

Figure 5.6. Fuzzy c-means clustering of ANOVA significant proteins.

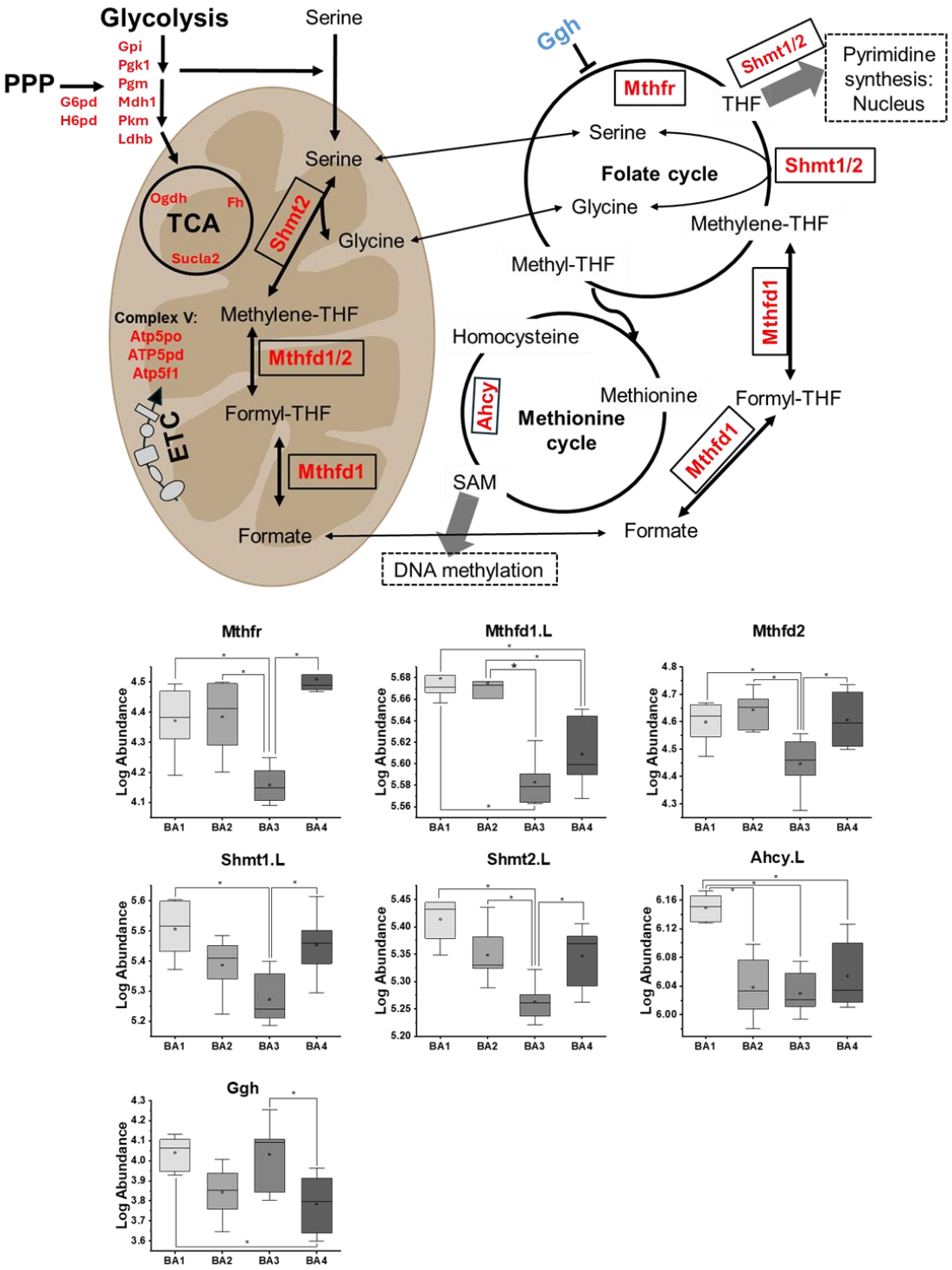

Figure 5.7. Schematic of the folate cycle and other connecting metabolic pathways. Pathways and proteins enriched in Cluster 1 (top panel). Abundances of proteins involved in the folate cycle within four BAs. Key: Protein symbols in red are upregulated while proteins in blue are

downregulated in BA1, BA2, and BA4. PPP: Pentose phosphate pathway. Full forms in text or glossary.

Cluster 1 revealed the enrichment of biological processes related to folic acid metabolism, TCA cycle, glucose metabolic processes, ATP synthesis, and ribonucleoprotein complex subunit organization. **Figure 5.7** depicts the proteins under metabolic and folic acid pathways, illustrating their interconnections. We found enzymes involved in folic acid metabolism and methionine cycle upregulated in BA1, BA2, and BA4 whereas downregulated in BA3. Additionally, the enzyme Ggh (gamma-glutamyl hydrolase) which hydrolyzes the polyglutamate tail on folates to reduce their affinity to folate cycle enzymes was upregulated in BA3. Previous reports have suggested a tight regulation between folate and methionine cycle in CNC development.[268] Elevated homocysteine levels, coupled with reduced methionine and folate levels, facilitate the migration of neural crest cells, while the converse trend promotes differentiation. Regulation of the folate cycle is also dependent on glycolysis and serine synthesis. Consequently, glycolytic and serine/glycine biosynthesis enzymes were also upregulated in BA1, BA2, and BA3 (**Fig 5.7**). Furthermore, we also found the Mthfr Ish probe showing the same pattern in BAs (recall **Fig. 5.4**). However, the role of folate and methionine cycle in the differentiation and migration of BAs along the AP axis has not been explored yet. The proteomic analysis opens an exciting avenue to further explore the role of folate metabolism and glycolysis in craniofacial development.

Cluster 2 displays proteins that are upregulated in the anterior arches (BA1 and BA2) compared to posterior arches (BA3 and BA4). We observed the upregulation of proteins (Gsn, Lmnb2, Akt2, and Parp1) regulating the FAS signaling pathway which is involved in programmed cell death. Studies show that Gsn and Akt expression positively or negatively controls FAS mediated

apoptosis.[269-270] Similarly, Parp1 is known to interact with Faf1 to regulate FAS signaling[271], although the expression of Parp1 was found to be significantly depleted in BA3, its expression was similar between BA1, BA2, and BA4, interestingly Faf1 was found to be significantly enriched in BA4 indicating differential regulation of FAS signaling between the four arches. Due to the paucity of literature regarding the involvement of FAS-mediated signaling in cranial neural crest development, there is a need for further investigations to elucidate its potential role in this process.

We also detected proteins with known canonical roles in neural crest migration and differentiation. For instance, we found differential enrichment of proteins regulating RNA splicing within the BAs in cluster 2. Cirbp (Cold-inducible RNA binding protein), a regulator of anterior-posterior patterning was found to be enriched in BA1.[272] While cluster 3 did not show enrichment of any specific GO category, cluster 4 was found to be enriched in proteins related to actin filament bundle assembly. Actin filaments are required to control the speed and direction of neural crest migration. For example, we found the protein Cald1 (caldesmon 1) significantly upregulated in BA4 compared to the other three arches. Caldesmon is known to be highly expressed in the pre-migratory and post-migratory CNCs to guide CNC migration. A previous study in *Xenopus* CNCs showed localization of *cald1* transcript (Ish) in pre-migratory (stage 15) and migrating CNCs (stage 20). Surprisingly, contrary to our proteomic data, their Ish visualization did not find *cald1* in branchial arches at stage 26 (post-migratory). However, their knockdown experiment using *cald1* morpholino revealed migration defects as the third and fourth streams, labeled by *twist* and *sox10*, were not well separated matching our proteomic finding.[273] Our spatial dataset unveils a novel deep proteomic coverage of the CNC streams,

presenting not only the proteomic expression profiles of established canonical markers but also uncovering numerous potential markers for branchial arch (BA) differentiation and migration.

Proteomic profiling of pre-migratory CNCs vs post-migratory CNCs. Apart from spatial differences, we sought to profile proteomic remodeling between pre-migratory and post-migratory CNCs to identify prominent shifts within the proteome. We performed standard bottom-up proteomics on stage 17 (pre) and stage 26 (post) CNC pools with in-vial yolk depletion and analyzed the samples using LC-MS^2 label-free quantification (see *Methods*). We identified ~4000 proteins between the two temporal stages. We performed fold change analysis on log normalized data. Proteins that showed a fold change of 2 between the two temporal CNC stages were then tested using the statistical overrepresentation test to identify preferential enrichment of specific GO categories. **Figure 5.8** compares the biological processes enriched in pre vs post-CNCs. One major shift observed involves the change in energy pathways. While pre-CNC shows upregulation of fatty acid synthesis and aerobic respiration (ETC), glucose metabolism and glycolytic processes emerged as significant metabolic pathways in post-CNCs. This shift is likely important to regulate the folate cycle for spatial differentiation of the BAs as discussed in the previous section.

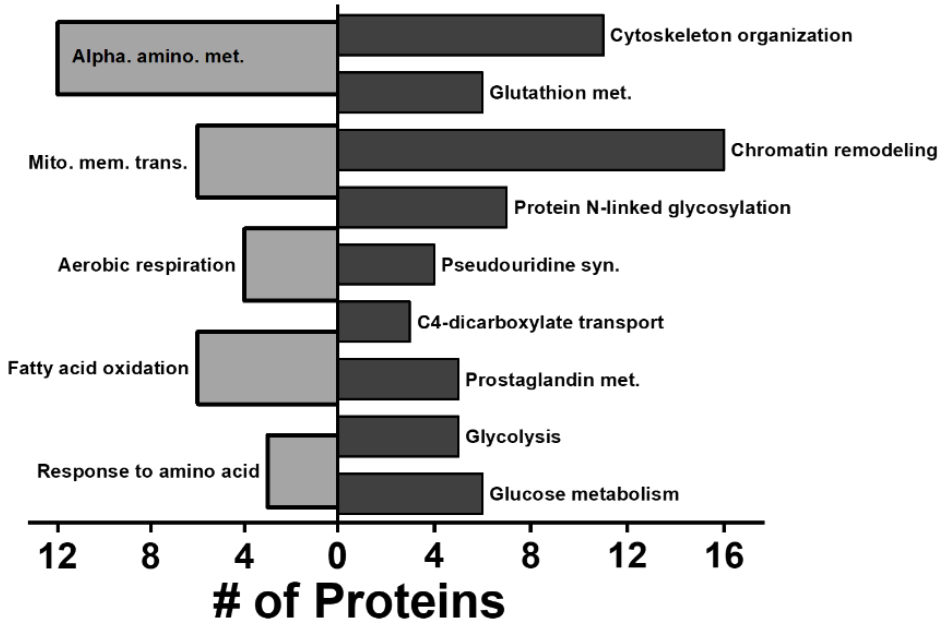

Figure 5.8. Comparison of biological processes enriched in pre-migratory CNCs vs post-migratory CNCs. Statistical enrichment test performed on proteins showing over 2-fold change in protein abundance. Key: Alpha. amino. met., alpha-amino acid metabolic process; Mito. mem. trans., mitochondrial transmembrane transport; glutathione met., glutathione metabolism; pseudouridine syn., pseudouridine synthesis; prostaglandin met., prostaglandin metabolism.

Other biological processes enriched in post-CNCs include chromatin remodeling, cytoskeletal filament organization, pseudouridine synthesis, protein N-linked glycosylation, prostaglandin metabolic process, glutathione metabolic process, and C4-dicarboxylate transport. Upregulation of chromatin remodeling indicates active transcription and translation of new genes as it makes chromatin strands accessible by allowing the binding of transcription factors and DNA binding proteins. Cytoskeletal reorganization is understandable as the pre-migratory CNCs slowly diverge into four streams. Notably, we also found upregulation of enzymes involved in pseudouridine synthesis, which is known to be important for the stability of RNA and to guide the splicing process, however, its specific role in the CNC development remains elusive. Furthermore, we observed an enrichment of glutathione metabolism and dicarboxylate carriers which are known to be important to maintain ROS homeostasis implying increased ROS in

migrating CNCs. Due to limited information on the role of these noncanonical biological processes in pre-migratory CNCs and BAs, our proteomic dataset has the potential to serve as a valuable resource for generating novel hypotheses and revealing previously undiscovered mechanisms.

5.5 Conclusions

Our study on deep spatial proteomics of BAs has provided valuable insights into the embryonic development process. Initially, we evaluated yolk content in pre-migratory and post-migratory cranial neural crest cells (CNCs), indicating the need for yolk depletion to enhance proteomic coverage in these cell populations. We devised and implemented the YODEC approach for diluting the yolk in BAs, as the limited population of CNC/BA cells (~1000-2000 cells) prevented the yolk depletion of individual samples. Utilizing this approach, we identified over 8000 proteins, with approximately 5000 proteins showing quantitative expression in multiple biological replicates. Unsupervised hierarchical clustering highlighted significant differences in protein expression across the four BAs, with distinct patterns observed, particularly in BA3. Comparative analysis with existing genomic expression data validated our findings, indicating potential roles for identified proteins in craniofacial development. Additionally, ANOVA analysis identified nearly 900 proteins differentially regulated among the BAs, with fuzzy c-means clustering categorizing them into distinct expression profiles. Enrichment analysis revealed the involvement of various biological processes, including folic acid metabolism, FAS signaling, and actin filament bundle assembly, suggesting their significance in BA differentiation and migration.

Based on our proteomic data we hypothesize that the folate and methionine cycle is critical for BA migration across the AP axis or BA differentiation. We identified numerous proteins as potential markers for spatial craniofacial development such as Mthfr, Ahcy, Denr, Rmbx, Cald1, and Gsn. We discovered the novel role of the folate cycle and its differential expression within the BAs. Moreover, the preliminary comparison between pre-migratory and post-migratory CNCs highlighted global metabolic shifts, particularly in energy pathways, such as aerobic respiration vs glycolysis. We posit that such metabolic shift plays a crucial role in regulating the folate cycle for BA spatial differentiation. How the folate cycle affects CNC/BA migration and differentiation will be the focus of our future studies. Overall, our comprehensive proteomic profiling offers novel insights into BA development, uncovering potential markers and mechanisms previously unexplored. This dataset not only advances our understanding of craniofacial morphogenesis but also sets the stage for future investigations into the intricate molecular mechanisms underlying embryonic development.

Acknowledgements. This work was funded by the University of Maryland-National Cancer Institute Partnership Program (seed award to P.N.) and Daar laboratory at the National Cancer Institute, NIH. We thank Annie Liu for assisting with protein name annotations of the proteomic dataset. We thank Dr. Jaeho Yoon for assisting with embryology, imaging, and collecting samples for proteomic analysis. We thank the Daar laboratory for providing resources for the proteomic analysis.

Chapter 6: Conclusions and Future Directions

The overarching goal of my research was to develop and integrate novel proteomic approaches to enable spatiotemporal analysis in limited embryonic tissues to answer biological questions about embryonic patterning during healthy development. The following section summarizes the analytical developments and novel biological insights discovered during the progression of my Ph.D. work.

6.1 Analytical developments

Here, we focused on improving the proteomic coverage and detection of proteins from embryonic cell populations. While significant progress has been made across all aspects of MS workflow, the challenge posed by highly abundant proteins impeding the detection of biologically significant low-abundance proteins remains unresolved. Especially in *Xenopus* tissues, we demonstrated the severity of yolk interference on proteomic coverage. ~90% of the total protein abundance was contributed by the yolk, necessitating the need for new approaches to address this challenge. As the macroscale yolk depletion method is not amenable to small cell populations (<5000 cells), we developed two analytical techniques, to dilute the yolk interference using an intelligent detection strategy and a sample preparation strategy to physically isolate yolk platelets from miniscule cell populations.

At first, we designed a detection strategy using isobaric mass tags to dilute the yolk proteins and preclude the need to individually yolk deplete limited biological samples. YODEC relied on a depleted carrier sample spiked at a higher amount in the biological samples. The depleted carrier diluted the yolk peptides and boosted the abundance of nonyolk peptides, which likely enabled their selection for tandem MS, allowing them to be sequenced. Furthermore, the isobaric mass

tags allowed relative quantification between biological samples of the newly identified proteins, thus boosting quantitative proteomic coverage. In parallel, we developed a microscale sample preparation approach (µYODE) to deplete yolk platelets from embryonic tissues reproducibly. Excessive dilution of tissues with sample buffer leads to sample losses caused by adsorption on vial surfaces. Therefore, the approach centered on devising a microscale setup to deplete yolk in a capillary format from limited volumes (< 10 µL) of cell lysate. Each of these developments doubled the protein quantification from < 500 ng of yolk free protein amounts (< 5 µL total protein amount) compared to the traditional strategies. Both approaches minimized the necessity for cumbersome sample pooling, consequently enabling us to augment the number of biological replicates for the tissues under investigation. This enhancement in replicate numbers bolstered the statistical confidence in our proteomic data.

Apart from the above mentioned tailored approaches to minimize yolk interference, we also tested other parameters of the LC-MS workflow. As global proteomic analysis encounters complex protein digests, usually a separation technique is coupled to MS to reduce sample complexity before MS analysis and maximize the duty cycle. We optimized our LC method using the new generation µPAC technology which reduces peak broadening and improves resolution. To further reduce sample complexity, we performed offline fractionation of our complex digest at high pH followed by low pH LC-MS analysis. The amphoteric nature of the peptides facilitates orthogonal separation at different pH levels. Additionally, we tested different emitters for the LC-ESI interface to improve the ionization of the peptide molecules. Integrating all these developments together afforded an 8-fold increase in protein identifications from limited embryonic tissues. Considerable effort was dedicated to refining the technology prior to

delving into the biological study, ensuring that we attained sufficient sensitivity to investigate the biological question.

6.2 Biological insights

Studying embryonic patterning in healthy embryos holds significant importance in systems biology, serving as a crucial avenue for comprehending disease progression, regeneration, and the development of novel therapies. Primarily, we focused on two important cell populations in early embryonic development, the organizer and the CNCs. The SMO is an important signaling center established at the beginning of gastrulation in the *Xenopus* model which guides neural induction and dorsoanterior development during gastrulation. We conducted spatial proteomics analyses on the SMO, induced NE, and entire embryo, aiming to reveal the spatial proteomic heterogeneity within the embryo that facilitates the progression of healthy gastrulation. The goal of the study was to reveal key proteomic gradients to identify novel markers for neural induction and dorsal development. Using our developed proteomic approach, we identified ~8000 proteins from single dissections (< 5000 cells) from these tissues. We uncovered differential enrichment of metabolic pathways between the SMO and the NE. This discovery further led us to discover the novel role of H_2O_2 signaling during gastrulation. We designed orthogonal studies to validate the role of localized H_2O_2 during gastrulation. By perturbing the H_2O_2 localization in the SMO, we demonstrated the defects in the dorsoanterior and neural structures in the *Xenopus* model, thereby confirming the pertinent role of the H_2O_2 gradient during gastrulation.

In the post gastrulation stages, the CNCs play an important role in craniofacial development. These multipotent progenitor cells are used as a model to understand stem cell differentiation and migratory behavior during cancer metastasis. During development the CNCs migrate, diverge,

and give rise to four streams called the BAs. The four BAs then differentiate into a multitude of craniofacial tissues. To understand the proteomic remodeling that allows this migratory behavior and differentiation we conducted a deep spatial proteomic analysis of the BAs. This is the first ever global proteomic dataset collected on the four branchial arches. We discovered the expression of proteins related to the methionine/folate cycle, glycolysis, FAS signaling, actin cytoskeletons differ between the four branchial arches. Utilizing mRNA probes based images available in the literature, we visually confirmed correlations between transcript and protein expression. The significant protein markers uncovered in this study will aid in the development of future investigations aimed at elucidating the functional roles of these molecules and generating new hypotheses. In summary, the technological innovations developed in this work can be adapted across different biological samples and species, while the comprehensive proteomic data has furnished novel insights that can lay the groundwork for future research endeavors aimed at enhancing our current understanding of embryonic development.

6.3 Future technological innovations

Advances in MS-based approaches continually enhance the achievable sensitivity for detecting proteins from various biological specimens. In this study, we tailored ultrasensitive proteomic techniques to investigate limited embryonic tissues. Although we achieved an 8-fold increase in coverage, the proteome remains considerably complex, and further developments in newer MS techniques could enhance sensitivity for detecting proteins with lower abundance. Emerging methods like DIA offer a potential solution to the sensitivity constraints associated with DDA, which tends to favor higher abundance. In contrast to DDA, DIA typically involves the selection of all ions within a specific m/z window, effectively capturing nearly all precursor ions regardless of their abundance.[274] Data collected via DIA are often complex and computationally

challenging to analyze. However, the advent of various AI-based strategies and user-friendly software interfaces has made data interpretation feasible.[275] While DIA acquisition has primarily been limited to label-free quantification, recent developments with non-isobaric mass tags suggest the potential for multiplexing with DIA.[276] Integrating YODEC or μYODE with DIA data acquisition could help pursue a near complete coverage of the proteome.

6.4 Leap towards future biological discoveries

We uncovered spatial proteomic gradients within the gastrulating embryo. Apart from canonical pathways, we discovered non-canonical metabolic differences between the SMO and the NE tissue. The NE tissue showed upregulation of fatty acid oxidation and TCA cycle enzymes, while the SMO tissue showed upregulation of glycolysis and pyruvate metabolism pathways. The role of differential metabolism within these spatially distinct lineages (neural vs. mesodermal) has not been explored. To validate the functional significance of these metabolic differences, the initial step involves examining the basal levels of these metabolites in the SMO and the NE. We attempted to directly profile the TCA and glycolysis metabolites from the neighboring tissues using GC-MS, however, due to sensitivity limitation we were only able to detect enrichment of pyruvate (glycolysis end product) in the SMO tissue. To pursue this investigation further, we need to develop a sensitive metabolomic assay capable of identifying all the metabolites from limited embryonic tissues. LC-MS analysis employing specialized ion exchange columns for separating small organic acids offers a promising alternative to achieve this goal as it eliminates the need for derivatization protocols required for GC-MS analysis.[277] Once a robust metabolomic method is established, the utilization of isotopic markers for metabolic flux analysis could further elucidate the rates of interconversion and the activity of these pathways. The metabolomic data

will help design future functional studies to perturb the pathways of interest and assess their effect on embryonic development.

Apart from metabolomic differences, we unearthed the novel role of the H_2O_2 gradient in dorsoanterior patterning. We revealed localized H_2O_2 expression in the SMO region. Using fluorescence imaging, we confirmed that H_2O_2 is generated at the beginning of gastrulation and remains localized in the SMO until the end of gastrulation. We further validated the time-sensitive expression of H_2O_2, which is important, and perturbing the gradient affects normal embryonic development through functional experiments. Nonetheless, the significance of H_2O_2 expression during gastrulation and its subsequent downstream interactions remain to be examined. The presence of H_2O_2 expression may suggest that downstream proteins undergo oxidation, indicating that studying oxidation PTMs could unveil potential markers. Post-translational modifications due to oxidation typically occur on methionine or cysteine residues of proteins.[278] Utilizing iodoTMT-based multiplexing kits can facilitate the labeling of cysteine residues in various biological samples, enabling the profiling of oxidation PTMs. Additionally, mapping the temporal proteomic patterns between the SMO and the NE from early stages (precursor cells at NF stage 32) to post-gastrulation will aid in understanding how the proteome evolves over time. This approach will pinpoint critical stages where the proteome of embryonic cells begins to diverge as they adopt distinct lineages. A deep proteomic dataset of gastrulation events, tissue induction, and differentiation would serve as a valuable resource for generating future hypotheses.

List of Publications

The work presented in this dissertation has been disseminated as follows:

1. **L. R. Pade,** C. Lombard-Banek, J. Li, and P. Nemes*, “Dilute to enrich for deeper proteomics: A yolk-depleted carrier for limited populations of embryonic (frog) cells” in: Journal of Proteome Research, 23 (2): 692–703, https://doi.org/10.1021/acs.jproteome.3c00541 **(cover page article)**

2. A. B. Baxi, J. Li, V. M. Quach**, L. R. Pade**, S. A. Moody, and P. Nemes*, “Cell-lineage-guided mass spectrometry reveals increased energy metabolism and reactive oxygen species in the vertebrate organizer” in **PNAS**, 121 (6), e2311625121 (2024)

3. **L. R. Pade**, C. Lombard-Banek, J. Li, and P. Nemes*, Dilute to enrich for deeper proteomics: A yolk-depleted carrier for limited populations of embryonic (frog) cells, ChemRxiv submitted on 8/21/2023, https://chemrxiv.org/engage/chemrxiv/article-details/64e3b00e00bbebf0e690dc6f

4. **L. R. Pade**, K. E. Stepler, E. P. Portero, K. DeLaney, and P. Nemes. “Mass spectrometry enables spatiotemporal ‘omics: from tissues to cells to organelles” in: Mass Spectrometry Reviews, 43 (1): 106–138, https://doi.org/10.1002/mas.21824 (2023).

5. E.P. Portero*, **L.R. Pade***, J. Li, S. Choi, P. Nemes, “Single-cell mass-spectrometry of metabolites and proteins for systems and functional biology” in: Single Cell ‘Omics of Neuronal Cells, edited by J. V. Sweedler, J. H. Eberwine, S. E. Fraser, Neuromethods Series, 184, 87–114, https://doi.org/10.1007/978-1-0716-2525-5_5 (2023).

6. B. Shen, **L. R. Pade**, S. B. Choi, P. Muñoz-LLancao, M. C. Manzini, and P. Nemes. “Capillary electrophoresis mass spectrometry for scalable single-cell proteomics” in: Frontiers in Chemistry, 1–11, https://doi.org/10.3389/fchem.2022.863979 (2022).

7. A.B. Baxi, **L.R. Pade**, P. Nemes. “Cell-lineage guided mass spectrometry proteomics in the developing (frog) embryo” in: J. Vis. Exp., e63586, doi:10.3791/63586 (2022).

8. A.B. Baxi, **L.R. Pade**, P. Nemes. "Mass spectrometry based proteomics for neurodevelopmental biology in the amphibian *Xenopus laevis*" in: Current topics in developmental biology, 145: 205-231, 10.1016/bs.ctdb.2021.04.002 (2021).

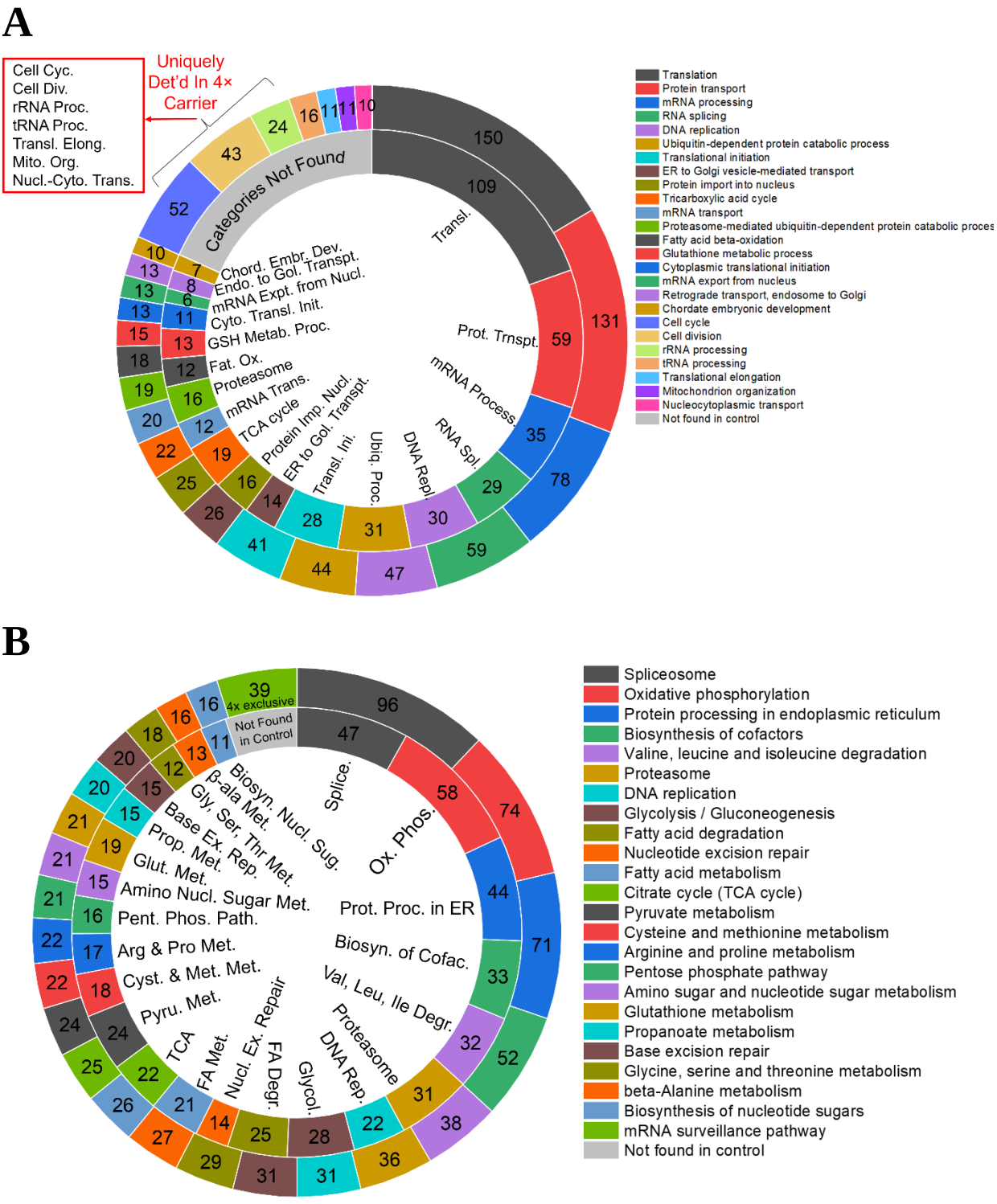

Figure 1 Continued…

C

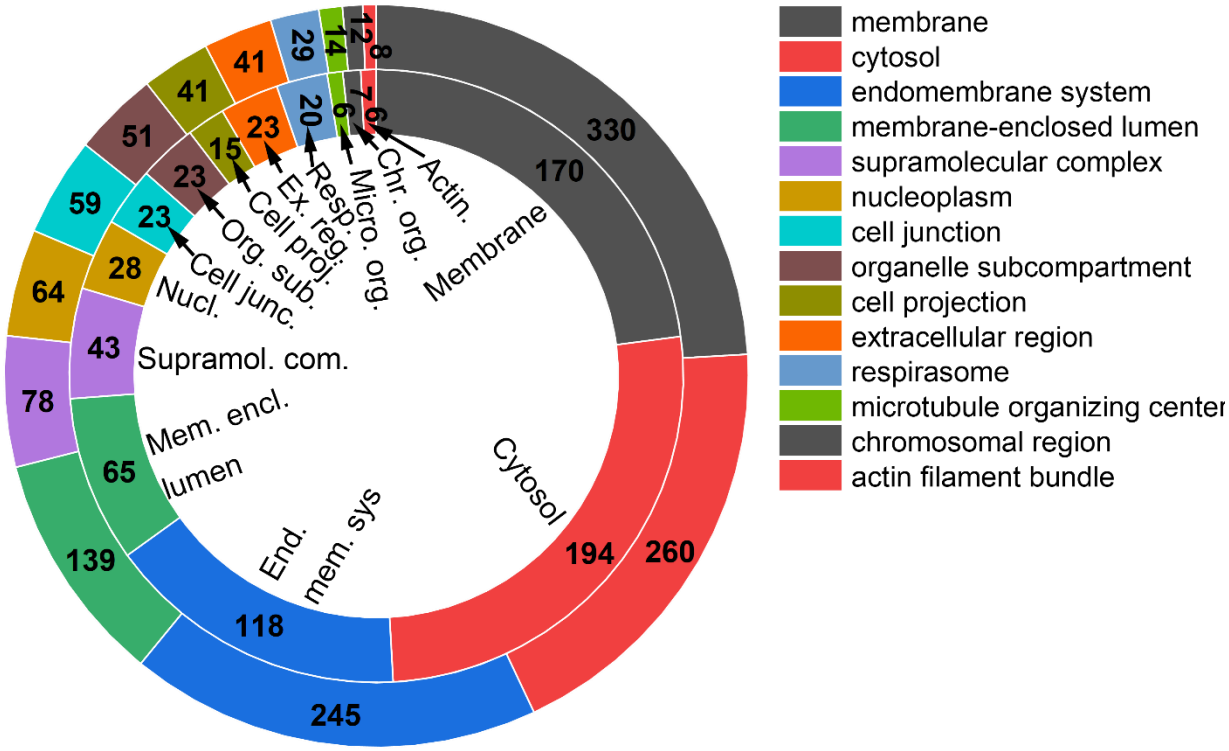

Figure A 1. GO analysis of biological processes, signaling pathways, and subcellular distribution. Compared to the control, YODEC improved protein coverage in (A) biological processes, (B) signaling pathways, and (C) subcellular distribution. The inner and outer circles map protein categories in the Control and YODEC (4-fold carrier load), respectively. Each GO category is color coded.

A

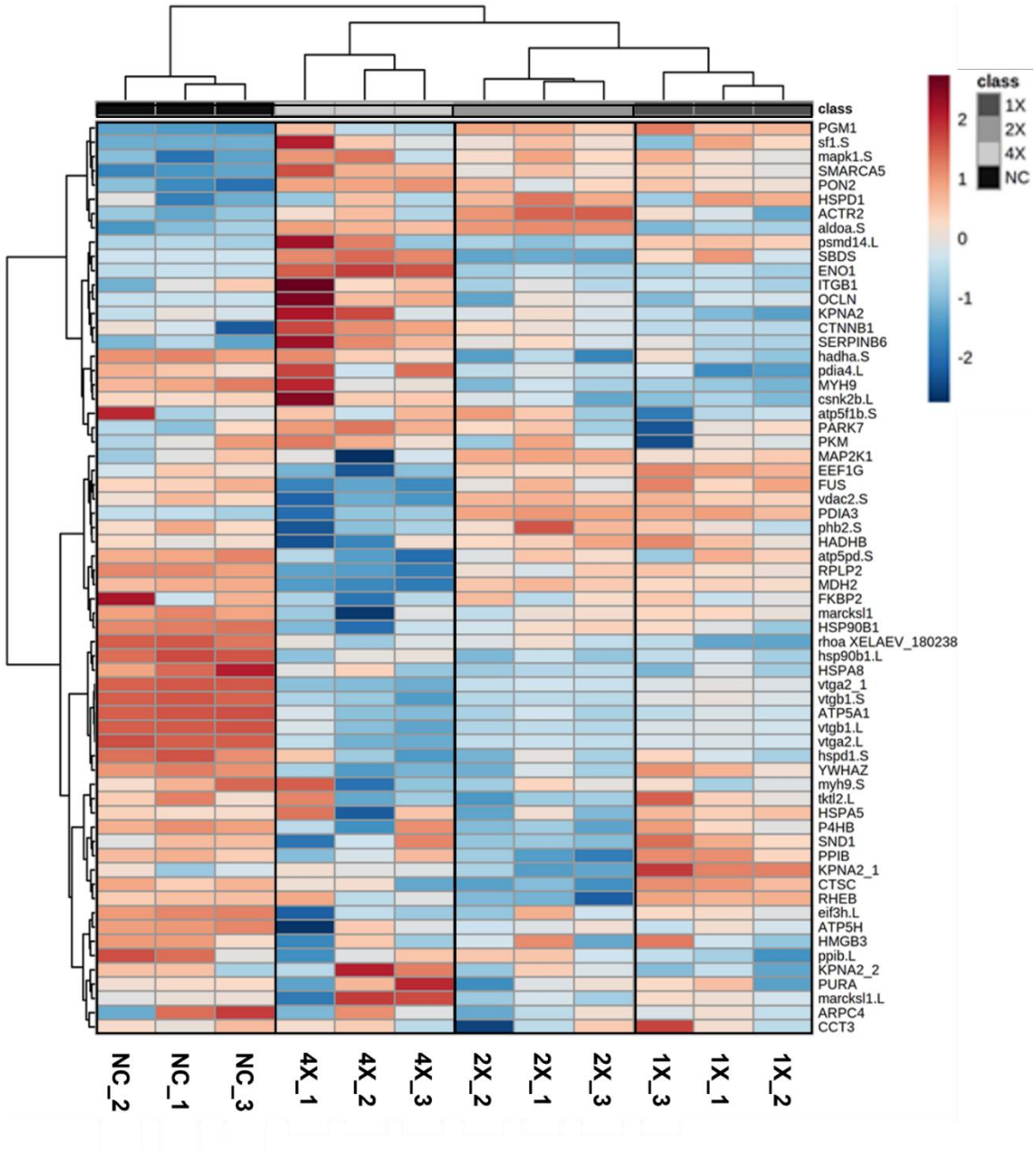

Figure A 2 Continued…

B

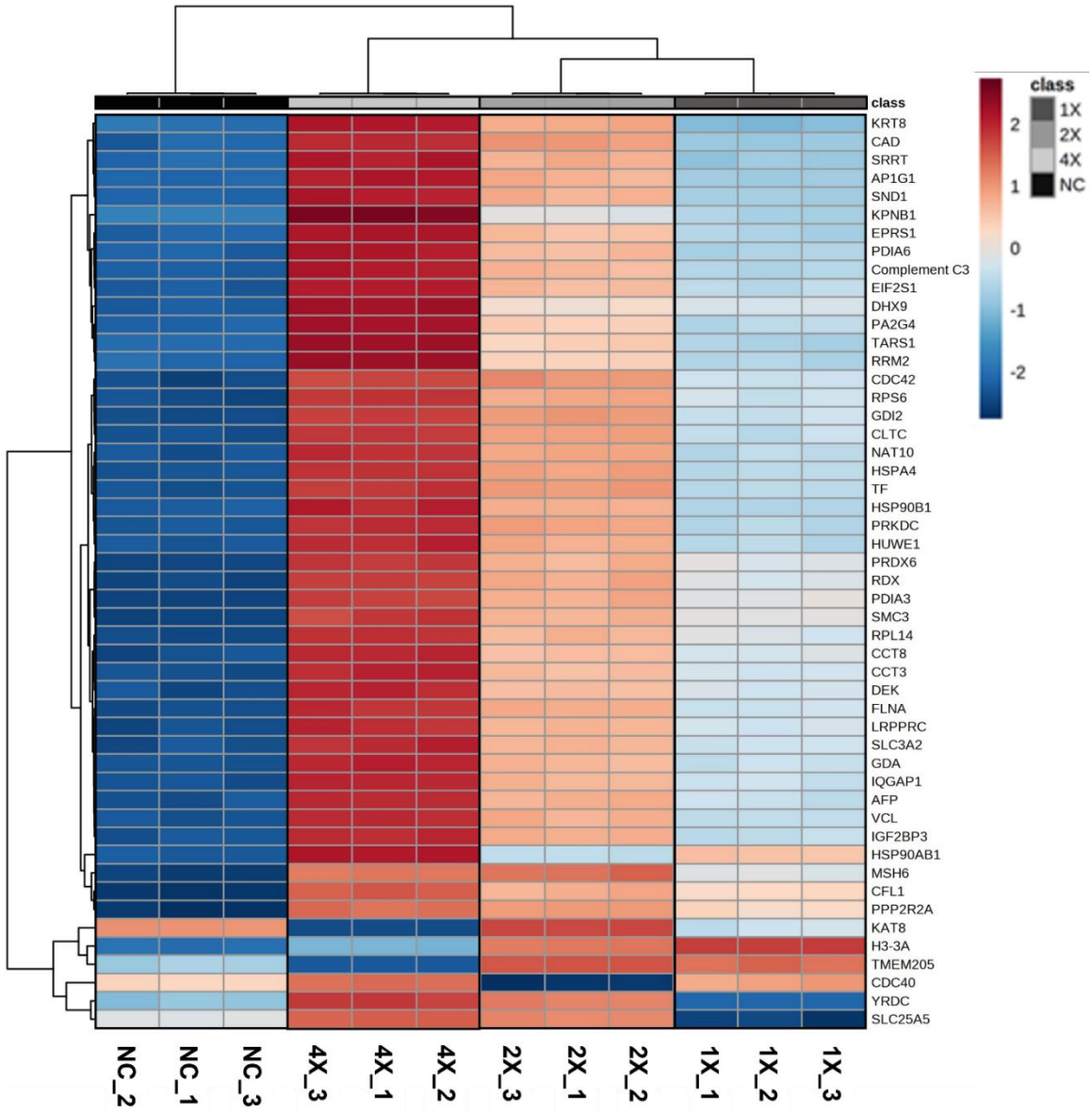

Figure A 2. Experimental validation of the proposed working principle by YODEC (A) Detailed heatmap of top 50 proteins diluted or enriched in the YODEC experiment with increasing deyolked carrier loads. (B) HeLa heatmap after serial protein dilution, revealing a lack of differential enrichment among the proteins.

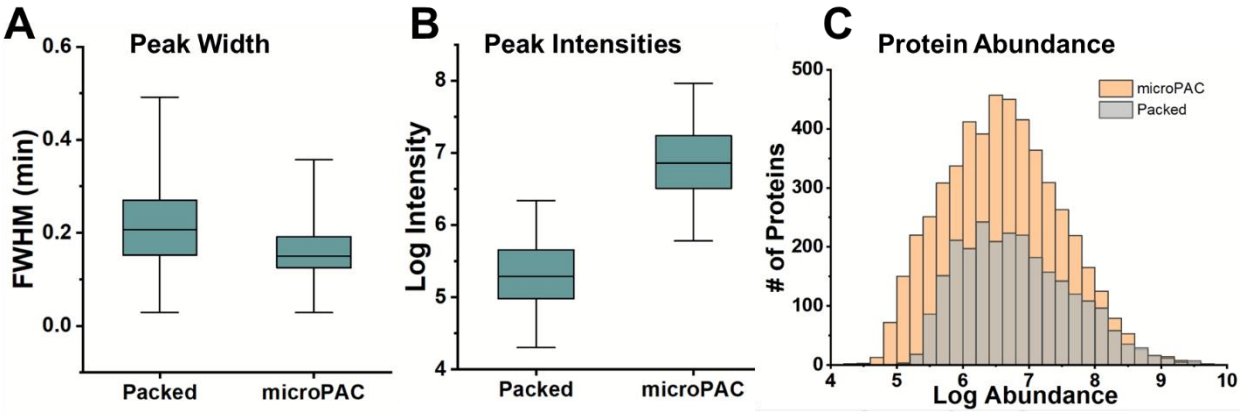

Figure A 3. Comparison of chromatographic parameters between packed bed column and µPAC. **(A)** Peak widths of precursor ions for commonly identified precursors show µPAC achieves narrower peaks improving resolution. **(B)** Peak intensity distribution of commonly detected precursor ions displaying higher intensity resulting from sharper peaks. **(C)** Protein abundance (log transformed) of all proteins detected by the packed bed and µPAC indicating several new proteins are detected in the lower abundance range.

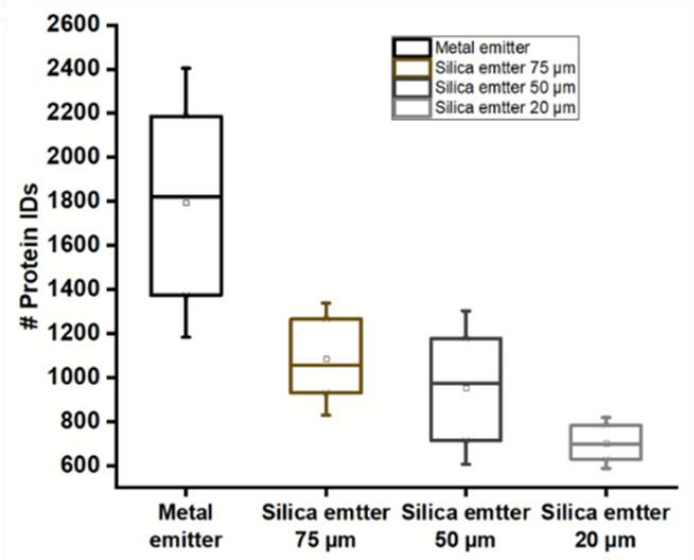

Figure A 4. Comparing nanoESI interface design to improve ionization. Metal emitter with a direct junction interface provided higher sensitivity and protein identifications compared to silica emitters with varying inner diameters in a liquid junction interface.

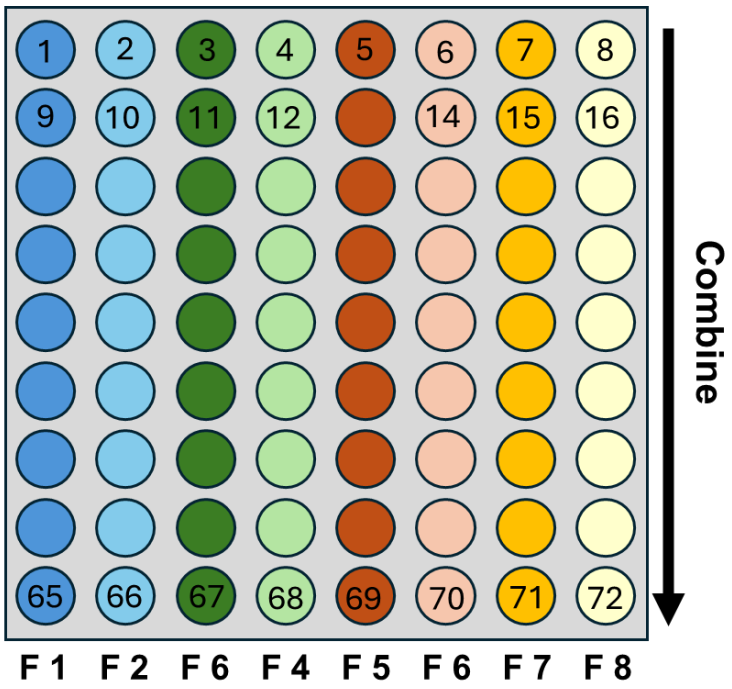

Figure A 5. Concatenation strategy to orthogonally combine 72 fractions to 8 fractions for LC-MS analysis. (Key: F, fraction no.)

Table A 1. Examples of proteomics by HRMS supporting neurodevelopmental studies in *Xenopus.* Reproduced from reference[71].

Targeted Cell/Tissue Type	Modes of Separation and MS	Type of Quantification	No. of Proteins Identified	Reference
Retina	2D DIGE - RPLC-MS	Label-free	~2,000	[279]
Otic vesicles	High-pH RPLC-Low-pH RPLC-MS	Relative (TMT)	~5,000	[280]
Spinal cord tissues	SCX-RPLC-MS	Relative (iTRAQ)	~6,000	[281]
Neural ectoderm tissues (beginning of gastrulation)	High-pH RPLC-Low-pH RPLC-MS	Label-free	~2,500	[42]
Spemann organizer tissue (beginning of gastrulation)	High-pH RPLC-Low-pH RPLC-MS	Relative (TMT)	~3,000	[282-283]
Axons	RPLC-MS	Metabolic labeling (pSILAC)	~350 labeled proteins	[284]
Axons	MudPIT-MS	Metabolic labeling (BONCAT)	~4,800	[285]
Neural, epidermal, and endodermal fated single cells	CE-MS	Label-free	~438	[286]
Neural fated single cells, visual assay of the tadpole	CE-MS	Label-free	~738	[27]

Table A 2. GO annotation (biological processes) for proteins displayed in the heatmap differentiating the SMO, the NE, and the WE tissue (**Fig. 4.4**). Columns are color-coded to represent the proteins enriched in each tissue type, NE: red; SMO: green; WE: gray

Heat map Accession	Protein Symbol (*X. laevis*)	Protein Symbol (Human)	Biological process
rna78556|	Rps24	Rps24	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA
CL6048Contig2	Psma4	Psma4	ubiquitin-dependent protein catabolic process
rna29490|	Gstm1.L	Gstm1	glutathione metabolic process
CL8328Contig1	Adss2	Adss	'de novo' AMP biosynthetic process
CL1741Contig4	Rps15	Rps15	ribonucleoprotein complex assembly
CL10410Contig1	Dhfr	Dhfr	folic acid-containing compound biosynthetic process
rna33584|	Psmb1.S	Psmb1	proteolysis involved in cellular protein catabolic process
CL1Contig1626	Eif2s1	Eif2s1	translational initiation
rna76002|	Capzb.S	Capzb	negative regulation of actin filament polymerization
rna18040|	Hadh.L	Hadh	fatty acid metabolic process
rna21502|	Blmh.L	Blmh	cysteine-type peptidase activity
CL2097Contig2	Ahcy	Ahcy	hydrolase activity
rna28152|	Ncbp1.L	Ncbp1	RNA cap binding
rna40817|	Pa2g4.S	Pa2g4	transcription
rna83639|	Asl.L	Asl	alpha-amino acid metabolic process
CL9881Contig1	Spg7	Spg7	proteolysis
CL5824Contig1	Ptgs1	Ptgs1	vasculogenesis
CL4112Contig4	Myl6	Myl6	calcium ion binding
CL9369Contig1	Prxl2b	Fam213b	oxidation-reduction process
rna42000|	Get4.S	Get4	protein localization to organelle
CL5037Contig2	Ubfd1	Ubfd1	cell-cell adhesion
CL215Contig8	Bzw1	Bzw1	transcription
CL9166Contig2	Hsd17b12	Hsd17b12	fatty acid biosynthetic process
CL3111Contig3	Ipo9	Ipo9	protein import into nucleus
rna89830|	Ddah1	Ddah1	regulation of lipid biosynthetic process
CL14502Contig1	Tstd3	Tstd3	thiosulfate sulfurtransferase
rna37940|	Pak2.L	Pak2	protein serine/threonine kinase activity
CL1Contig787	Capn2	Capn2	proteolysis
CL9791Contig1	Ecpas	Ecm29	proteasome adaptor
CL3649Contig1	Cyb5r3	Cyb5r3	cholesterol biosynthetic process
CL384Contig7	Rpl7a	Rpl7a	RNA binding
rna30580|	Gins1.S	Gins1	mitotic DNA replication
rna75496|	Rtn3.L	Rtn3	vesicle-mediated transport

CL5228Contig1	Cops6	Cops6	protein deneddylation
CL159Contig1	Prmt1	Prmt1	transcription
rna42162|	Rab3a.S	Rab3a	organelle localization
CL4450Contig1	Rps6ka3	Rps6ka3	protein serine/threonine kinase activity
CL1282Contig7	Ap1g1	Ap1g1	intracellular protein transport
CL11588Contig1	Ufl1	Ufl1	ubiquitin-like protein transferase activity
CL4759Contig1	Psmd13	Psmd13	proteasome assembly
CL12697Contig1	Dnph1	Dnph1	nucleoside phosphate catabolic process
rna49089|	Ndufs6.S	Ndufs6	mitochondrial electron transport
CL303Contig4	Flna	Flna	regulation of cell migration
CL3725Contig3	Clns1a	Clns1a	spliceosomal snRNP assembly
rna38751|	Rabif.S	Rabif	post-Golgi vesicle-mediated transport
CL1Contig1307	Ndufs4	Ndufs4	electron transport chain
rna47136|	Crp.2	Crp	structural molecule activity
gene2971	Serpina3k.L	Serpina3k	serine-type endopeptidase inhibitor activity
CL8653Contig1	Itln2	Itln2	carbohydrate binding

Glossary

Aldoa	Aldolase, fructose-bisphosphate A
Tp1	Triosephosphate isomerase 1
Gapdh	Glyceraldehyde-3-phosphate dehydrogenase
Pgam5	Mitochondrial Serine/threonine-protein phosphatase
Eno1	Alpha-enolase
Pkm	Pyruvate kinase M
Ldha	Lactate dehydrogenase-A
Pdk1	Pyruvate dehydrogenase kinase 1
Mdh1	Malate dehydrogenase 1
Got1	Glutamic-oxaloacetic transaminase 1
Acat2	Acetyl-Coenzyme A acetyltransferase 2
Cpt2	Carnitine palmitoyltransferase 2
Aco2	Aconitase 2
Cs	Citrate synthase
Mdh2	Mitochondrial malate dehydrogenase
Fh	Fumarate hydratase
Sdhb	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial
Suclg2	GTP-specific b subunit of succinyl-coa synthetase
Dhtkd1	Dehydrogenase E1 and transketolase domain-containing protein 1
Ogdh	Oxoglutarate dehydrogenase
L2hgdh	L-2-hydroxyglutarate dehydrogenase
Idh3b	Isocitrate dehydrogenase (NAD(+)) 3 non-catalytic subunit beta
Ahcy	Adenosylhomocysteinase
Atp5f1	ATP synthase subunit b, mitochondrial
Atp5pd	Atp synthase peripheral stalk subunit d
Atp5po	ATP synthase peripheral stalk subunit OSCP
Fh	Fumarate hydratase
G6pd	Glucose-6-phosphate dehydrogenase
Ggh	Gamma-glutamyl hydrolase
Gpi	Glucose phosphate isomerase
H6pd	Hexose-6-phosphate dehydrogenase/glucose 1-dehydrogenase
Ldhb	Lactate dehydrogenase-B
Mdh1	Malate dehydrogenase 1
Mthfd1	Methylenetetrahydrofolate dehydrogenase 1/2
Mthfd1/2	Methylenetetrahydrofolate dehydrogenase 1/2
Mthfr	Methylenetetrahydrofolate reductase
Ogdh	2-oxoglutarate dehydrogenase E1 component

Pgk1	Phosphoglycerate kinase
Pgm	Phosphoglucomutase
Pkm	Pyruvate kinase isozymes M1/M2
Shmt1/2	Serine hydroxymethyltransferase 2 [(human)] 1/2
Shmt2	Serine hydroxymethyltransferase 2 [(human)]
Sucla2	Succinyl-coa ligase [ADP-forming] subunit beta, mitochondrial
Psmb1	Proteasome subunit beta type-1
Psma4	Proteasome subunit alpha type-4
Psmd13	26S proteasome non-atpase regulatory subunit 13
Ubfd1	Ubiquitin family domain containing 1
Ecpas	Ecm29 proteasome adaptor and scaffold
Pa2g4	Proliferation-associated 2g4
Hadh	Hydroxyacyl-coa Dehydrogenase
Cops6	Cop9 signalosome subunit 6
Asl	Argininosuccinate lyase
Rps15	Ribosomal protein s15
Rpl7a	Ribosomal Protein l7a
Eif2s1	Eukaryotic translation initiation factor 2 subunit alpha
Rps24	Ribosomal protein s24
Vim	Vimentin
Fn1	Fibronectin 1
Gsn	Gelsolin
Rps20	Ribosomal protein s20
Ugp2	Udp-glucose pyrophosphorylase 2
Pygl	Glycogen phosphorylase l
Tubal3	Tubulin alpha like 3
Cct6a	Chaperonin containing tcp1 subunit 6a
Stip1	Stress induced phosphoprotein 1
Polr2a	Rna polymerase ii subunit a
Mthfd1l	Methylenetetrahydrofolate dehydrogenase (nadp+ dependent) 1 like
Mthfr	Methylenetetrahydrofolate reductase
Ahcy	Adenosylhomocysteinase
Ipo7	Importin7
Ndufs5	Nadh:ubiquinone oxidoreductase subunit s5
Kpnb1	Karyopherin subunit beta 1
Nup54	Nucleoporin 54
Immt	Inner membrane mitochondrial protein
Atp5f1b	Atp synthase f1 subunit beta
Atp5pd	Atp synthase peripheral stalk subunit d
C3	Complement C3

Mpo	Myeloperoxidase
Ddx17	Dead-box helicase 17
Rbmx	Rna binding motif protein x-linked
Rbm15	Rna binding motif protein 15
Aars1	Alanyl-trna synthetase 1
Aimp2	Aminoacyl trna synthetase complex interacting multifunctional protein 2
Iars1	Isoleucyl-trna synthetase 1
Kars1	Lysyl-trna synthetase 1
Qars1	Glutaminyl-trna synthetase 1
Pkm	Pyruvate kinase m1/2
Rps2	Ribosomal protein s2
Rpl26	Ribosomal protein l26
Polr3f	Rna polymerase iii subunit f
Polr2b	Rna polymerase ii subunit b
Polr2a	Rna polymerase ii subunit a
Rpl27	Ribosomal protein l27
Rps7	Ribosomal protein s7
Rps23	Ribosomal protein s23
Eif1b	Eukaryotic translation initiation factor 1b
Denr	Density regulated re-initiation and release factor
Glu	Glucose
Glu-6-P	Glucose-6-phosphate
Fruc-6-P	Fructose-6-phosphate
Fruc-1,6-P	Fructose-1,6-bisphosphate
Glyc-3-P	Glyceraldehyde-3-phosphate
1,3-Diphosphogly	1,3-Diphosphoglycerate
3-Phosphogly	3-Phosphoglycerate
2-Phosphogly	2-Phosphoglycerate
Phosphoenolpyr	Phosphoenolpyruvate
IC	Isocitrate
KG	Alpha ketoglutarate
S-CoA	Succinyl CoA
OAA	Oxaloacetate
THF	Tetrahydrofolate
SAM	S-adenosylmethionine

Bibliography

1. De Robertis, E. M.; Larraín, J.; Oelgeschläger, M.; Wessely, O., The establishment of spemann's organizer and patterning of the vertebrate embryo. *Nat. Rev. Genet.* **2000,** *1* (3), 171-181.
2. Sato, T. S.; Handa, A.; Priya, S.; Watal, P.; Becker, R. M.; Sato, Y., Neurocristopathies: Enigmatic Appearances of Neural Crest Cell–derived Abnormalities. *RadioGraphics* **2019,** *39* (7), 2085-2102.
3. Cordero, D. R.; Brugmann, S.; Chu, Y.; Bajpai, R.; Jame, M.; Helms, J. A., Cranial neural crest cells on the move: their roles in craniofacial development. *Am J Med Genet A* **2011,** *155A* (2), 270-279.
4. Wessely, O.; Kim, J. I.; Geissert, D.; Tran, U.; De Robertis, E. M., Analysis of Spemann organizer formation in *Xenopus* embryos by cDNA macroarrays. *Dev. Biol.* **2004,** *269* (2), 552-566.
5. Gammill, L. S.; Bronner-Fraser, M., Neural crest specification: Migrating into genomics. *Nat. Rev. Neurosci.* **2003,** *4* (10), 795-805.
6. Suzawa, T.; Yoshida, H.; Takahashi, M.; Itose, M.; Takimoto, R.; Sasama, Y.; Tanaka, M.; Ikezaki, K.; Shirora, T.; Maki, K.; Kamijo, R., Prospects of neural crest-derived cells from oral and dentofacial tissues for application in regenerative medicine. *Oral Sci. Int.* **2020,** *17* (3), 115-125.
7. Sater, A. K.; Moody, S. A., Using *Xenopus* to understand human disease and developmental disorders. *Genesis* **2017,** *55* (1-2), 14.
8. Moody, S. A., Fates of the blastomeres of the 16-cell stage *Xenopus* embryo. *Dev. Biol.* **1987,** *119* (2), 560-578.
9. Moody, S. A., Fates of the blastomeres of the 32-cell stage *Xenopus* embryo. *Dev. Biol.* **1987,** *122* (2), 300-319.
10. Moody, S. A.; Kline, M. J., Segregation of fate during cleavage of frog (*Xenopus laevis*) blastomeres. *Anat. Embryol.* **1990,** *182* (4), 347-362.
11. Dale, L.; Slack, J. M., Fate map for the 32-cell stage of Xenopus laevis. *Development* **1987,** *99* (4), 527-551.
12. 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.
13. Perez, J. D.; Dieck, S. T.; Alvarez-Castelao, B.; Tushev, G.; Chan, I. C. W.; Schuman, E. M., Subcellular sequencing of single neurons reveals the dendritic transcriptome of GABAergic interneurons. *eLife* **2021,** *10*, 26.
14. Liu, Y. S.; Beyer, A.; Aebersold, R., On the dependency of cellular protein levels on mRNA abundance. *Cell* **2016,** *165* (3), 535-550.
15. Vogel, C.; Marcotte, E. M., Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **2012,** *13* (4), 227-232.
16. Peshkin, L.; Wuhr, M.; Pearl, E.; Haas, W.; Freeman, R. M.; Gerhart, J. C.; Klein, A. M.; Horb, M.; Gygi, S. P.; Kirschner, M. W., On the relationship of protein and mRNA dynamics in vertebrate embryonic development. *Dev. Cell* **2015,** *35* (3), 383-394.
17. 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.
18. Saadatpour, A.; Lai, S. J.; Guo, G. J.; Yuan, G. C., Single-cell analysis in cancer genomics. *Trends Genet.* **2015,** *31* (10), 576-586.
19. Tian, Q.; Stepaniants, S. B.; Mao, M.; Weng, L.; Feetham, M. C.; Doyle, M. J.; Yi, E. C.; Dai, H. Y.; Thorsson, V.; Eng, J.; Goodlett, D.; Berger, J. P.; Gunter, B.; Linseley, P. S.; Stoughton, R. B.; Aebersold, R.; Collins, S. J.; Hanlon, W. A.; Hood, L. E., Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol. Cell. Proteomics* **2004,** *3* (10), 960-969.
20. Feng, X. J.; Liu, X.; Luo, Q. M.; Liu, B. F., Mass spectrometry in systems biology: an overview. *Mass Spectrom. Rev.* **2008,** *27* (6), 635-660.
21. Geiger, T.; Wehner, A.; Schaab, C.; Cox, J.; Mann, M., Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol. Cell. Proteomics* **2012,** *11* (3), 1-11.
22. Hebert, A. S.; Richards, A. L.; Bailey, D. J.; Ulbrich, A.; Coughlin, E. E.; Westphall, M. S.; Coon, J. J., The one hour yeast proteome. *Mol. Cell. Proteomics* **2014,** *13* (1), 339-347.
23. Cervenka, J.; Tyleckova, J.; Skalnikova, H. K.; Kepkova, K. V.; Poliakh, I.; Valekova, I.; Pfeiferova, L.; Kolar, M.; Vaskovicova, M.; Pankova, T.; Vodicka, P., Proteomic characterization of human neural stem cells and their secretome during *in vitro* differentiation. *Front. Cell. Neurosci.* **2021,** *14*, 20.
24. Sperber, H.; Mathieu, J.; Wang, Y. L.; Ferreccio, A.; Hesson, J.; Xu, Z. J.; Fischer, K. A.; Devi, A.; Detraux, D.; Gu, H. W.; Battle, S. L.; Showalter, M.; Valensisi, C.; Bielas, J. H.; Ericson, N. G.; Margaretha, L.; Robitaille, A. M.; Margineantu, D.; Fiehn, O.; Hockenbery, D.; Blau, C. A.; Raftery, D.; Margolin, A. A.; Hawldns, R. D.; Moon, R. T.; Ware, C. B.; Ruohola-Baker, H., The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nat. Cell Biol.* **2015,** *17* (12), 1523-1535.
25. Arnold, P. K.; Jackson, B. T.; Paras, K. I.; Brunner, J. S.; Hart, M. L.; Newsom, O. J.; Alibeckoff, S. P.; Endress, J.; Drill, E.; Sullivan, L. B.; Finley, L. W. S., A non-canonical tricarboxylic acid cycle underlies cellular identity. *Nature* **2022,** *603* (7901), 477-481.
26. Gnann, C.; Cesnik, A. J.; Lundberg, E., Illuminating Non-genetic Cellular Heterogeneity with Imaging-Based Spatial Proteomics. *Trends in Cancer* **2021,** *7* (4), 278-282.
27. 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. Ed.* **2021,** *60* (23), 12852-12858.
28. Simone, N. L.; Bonner, R. F.; Gillespie, J. W.; Emmert-Buck, M. R.; Liotta, L. A., Laser-capture microdissection: opening the microscopic frontier to molecular analysis. *Trends Genet.* **1998,** *14* (7), 272-276.
29. Moroz, L. L.; Kohn, A. B., Single-neuron transcriptome and methylome sequencing for epigenomic analysis of aging. In *Biological Aging: Methods and Protocols, 2nd Edition*, Tollefsbol, T. O., Ed. Humana Press Inc: Totowa, 2013; Vol. 1048, pp 323-352.
30. OuYang, C. Z.; Liang, Z. D.; Li, L. J., Mass spectrometric analysis of spatio-temporal dynamics of crustacean neuropeptides. *BBA-Proteins Proteomics* **2015,** *1854* (7), 798-811.

31. Brown, H. M.; Pirro, V.; Cooks, R. G., From DESI to the MasSpec Pen: ambient ionization mass spectrometry for tissue analysis and intrasurgical cancer diagnosis. *Clin. Chem.* **2018,** *64* (4), 628-630.
32. Espino, J. A.; Jones, L. M., Illuminating biological interactions with *in vivo* protein footprinting. *Anal. Chem.* **2019,** *91* (10), 6577-6584.
33. 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.
34. von Eggeling, F.; Hoffmann, F., Microdissection-an essential prerequisite for spatial cancer omics. *Proteomics* **2020,** *20* (17-18), 10.
35. Yuan, Z. Y.; Zhou, Q. M.; Cai, L. S.; Pan, L.; Sun, W. L.; Qumu, S. W.; Yu, S.; Feng, J. X.; Zhao, H. S.; Zheng, Y. C.; Shi, M. L.; Li, S.; Chen, Y.; Zhang, X. R.; Zhang, M. Q., SEAM is a spatial single nuclear metabolomics method for dissecting tissue microenvironment. *Nat. Methods* **2021,** *18* (10), 1223-1232.
36. Azizian, N. G.; Sullivan, D. K.; Nie, L. T.; Pardo, S.; Molleur, D.; Chen, J. J.; Weintraub, S. T.; Li, Y. L., Selective labeling and identification of the tumor cell proteome of pancreatic cancer *in vivo*. *J. Proteome Res.* **2021,** *20* (1), 858-866.
37. Choi, S. B.; Polter, A. M.; Nemes, P., Patch-clamp proteomics of single neurons in tissue using electrophysiology and subcellular capillary electrophoresis mass spectrometry. *Anal. Chem.* **2022,** *94* (3), 1637-1644.
38. Arnadottir, S. S.; Mattesen, T. B.; Vang, S.; Madsen, M. R.; Madsen, A. H.; Birkbak, N. J.; Bramsen, J. B.; Andersen, C. L., Transcriptomic and proteomic intra-tumor heterogeneity of colorectal cancer varies depending on tumor location within the colorectum. *PLoS One* **2020,** *15* (12), 18.
39. Zhou, X. H.; Xiao, W.; Su, Z. Y.; Cheng, J. H.; Zheng, C. Y.; Zhang, Z. J.; Wang, Y. Q.; Wang, L.; Xu, B. H.; Li, S. P.; Yang, X. F.; Hoi, M. P. M., Hippocampal proteomic alteration in triple transgenic mouse model of Alzheimer's disease and implication of PINK 1 regulation in donepezil treatment. *J. Proteome Res.* **2019,** *18* (4), 1542-1552.
40. 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.
41. Zhang, Z. B.; Dubiak, K. M.; Shishkova, E.; Huber, P. W.; Coon, J. J.; Dovichi, N. J., High-throughput, comprehensive single-cell proteomic analysis of *Xenopus laevis* embryos at the 50-cell stage using a microplate-based MICROFASP System. *Anal. Chem.* **2022,** *94* (7), 3254-3259.
42. 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.
43. Hanson, J. N.; Motala, M. J.; Heien, M. L.; Gillette, M.; Sweedler, J.; Nuzzo, R. G., Textural guidance cues for controlling process outgrowth of mammalian neurons. *Lab Chip* **2009,** *9* (1), 122-131.
44. Hu, M. Z.; Helfenbein, K.; Buchberger, A. R.; DeLaney, K.; Liu, Y.; Li, L. J., Exploring the sexual dimorphism of crustacean neuropeptide expression using callinectes sapidus as a model organism. *J. Proteome Res.* **2021,** *20* (5), 2739-2750.

45. Shen, S. C.; Li, J.; Huo, S. A.; Ma, M.; Zhu, X. Y.; Rasam, S.; Duan, X. T.; Qu, M.; Titus, M. A.; Qu, J., Parallel, high-quality proteomic and targeted metabolomic quantification using laser capture microdissected tissues. *Anal. Chem.* **2021,** *93* (25), 8711-8718.
46. Zhu, Y.; Dou, M. W.; Piehowski, P. D.; Liang, Y. R.; Wang, F. J.; Chu, R. K.; Chrisler, W. B.; Smith, J. N.; Schwarz, K. C.; Shen, Y. F.; Shukla, A. K.; Moore, R. J.; Smith, R. D.; Qian, W. J.; Kelly, R. T., Spatially resolved proteome mapping of laser capture microdissected tissue with automated sample transfer to nanodroplets. *Mol. Cell. Proteomics* **2018,** *17* (9), 1864-1874.
47. Chappell, J.; Boward, B.; Dalton, S., Expanding the utility of FUCCI reporters using FACS-based 'omics analysis. In *Embryonic Stem Cell Protocols*, Turksen, K., Ed. Humana Press Inc: Totowa, 2016; Vol. 1341, pp 101-110.
48. Sergent-Tanguy, S.; Chagneau, C.; Neveu, I.; Naveilhan, P., Fluorescent activated cell sorting (FACS): a rapid and reliable method to estimate the number of neurons in a mixed population. *J. Neurosci. Meth.* **2003,** *129* (1), 73-79.
49. Chen, L.; Yang, R. R.; Kwan, T.; Tang, C.; Watt, S.; Zhang, Y. M.; Bourque, G.; Ge, B.; Downes, K.; Frontini, M.; Ouwehand, W. H.; Lin, J. W.; Soranzo, N.; Pastinen, T.; Chen, L., Paired rRNA-depleted and polyA-selected RNA sequencing data and supporting multi-omics data from human T cells. *Sci. Data* **2020,** *7* (1), 7.
50. Davey, H. M.; Kell, D. B., Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiol. Rev.* **1996,** *60* (4), 641-696.
51. Tharkeshwar, A. K.; Gevaert, K.; Annaert, W., Organellar omics-a reviving strategy to untangle the biomolecular complexity of the cell. *Proteomics* **2018,** *18* (5-6), 12.
52. Geladaki, A.; Britovsek, N. K.; Breckels, L. M.; Smith, T. S.; Vennard, O. L.; Mulvey, C. M.; Crook, O. M.; Gatto, L.; Lilley, K. S., Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics. *Nat. Commun.* **2019,** *10*, 15.
53. Landgraf, P.; Antileo, E. R.; Schuman, E. M.; Dieterich, D. C., BONCAT: metabolic labeling, click chemistry, and affinity purification of newly synthesized proteomes. *Methods Mol. Biol.* **2015,** *1266*, 199-215.
54. Schmerberg, C. M.; Li, L. J., Mass spectrometric detection of neuropeptides using affinity-enhanced microdialysis with antibody-coated magnetic nanoparticles. *Anal. Chem.* **2013,** *85* (2), 915-922.
55. Schmerberg, C. M.; Liang, Z. D.; Li, L. J., Data-independent MS/MS quantification of neuropeptides for determination of putative feeding-related neurohormones in microdialysate. *ACS Chem. Neurosci.* **2015,** *6* (1), 174-180.
56. Zestos, A. G.; Kennedy, R. T., Microdialysis coupled with LC-MS/MS for *in vivo* neurochemical monitoring. *Aaps J.* **2017,** *19* (5), 1284-1293.
57. Gupta, K.; Vuckovic, I.; Zhang, S.; Xiong, Y. N.; Carlson, B. L.; Jacobs, J.; Olson, I.; Petterson, X. M.; Macura, S. I.; Sarkaria, J.; Burns, T. C., Radiation induced metabolic alterations associate with tumor aggressiveness and poor outcome in glioblastoma. *Front. Oncol.* **2020,** *10*, 19.
58. Helbok, R.; Kofler, M.; Schiefecker, A. J.; Gaasch, M.; Rass, V.; Pfausler, B.; Beer, R.; Schmutzhard, E., Clinical use of cerebral microdialysis in patients with aneurysmal subarachnoid hemorrhage-state of the art. *Front. Neurol.* **2017,** *8*, 25.

59. Zhang, J. L.; Rector, J.; Lin, J. Q.; Young, J. H.; Sans, M.; Katta, N.; Giese, N.; Yu, W. D.; Nagi, C.; Suliburk, J.; Liu, J. S.; Bensussan, A.; DeHoog, R. J.; Garza, K. Y.; Ludolph, B.; Sorace, A. G.; Syed, A.; Zahedivash, A.; Milner, T. E.; Eberlin, L. S., Nondestructive tissue analysis for *ex vivo* and *in vivo* cancer diagnosis using a handheld mass spectrometry system. *Sci. Transl. Med.* **2017,** *9* (406), 11.
60. Fujii, T.; Matsuda, S.; Tejedor, M. L.; Esaki, T.; Sakane, I.; Mizuno, H.; Tsuyama, N.; Masujima, T., Direct metabolomics for plant cells by live single-cell mass spectrometry. *Nat. Protoc.* **2015,** *10* (9), 1445-1456.
61. Tsuyama, N.; Mizuno, H.; Tokunaga, E.; Masujima, T., Live single-cell molecular analysis by video-mass spectrometry. *Anal. Sci.* **2008,** *24* (5), 559-561.
62. Liu, R. M.; Pan, N.; Zhu, Y. L.; Yang, Z. B., T-Probe: an integrated microscale device for online *in situ* single cell analysis and metabolic profiling using mass spectrometry. *Anal. Chem.* **2018,** *90* (18), 11078-11085.
63. Taylor, M. J.; Mattson, S.; Liyu, A.; Stopka, S. A.; Ibrahim, Y. M.; Vertes, A.; Anderton, C. R., Optical microscopy-guided laser ablation electrospray ionization ion mobility mass spectrometry: ambient single cell metabolomics with increased confidence in molecular identification. *Metabolites* **2021,** *11* (4), 15.
64. 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.
65. 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.
66. 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.
67. Ctortecka, C.; Mechtler, K., The rise of single-cell proteomics. *Anal. Sci. Adv.* **2021,** *2* (3-4), 84-94.
68. Zhang, L.; Vertes, A., Single-cell mass spectrometry approaches to explore cellular heterogeneity. *Angew. Chem., Int. Ed.* **2018,** *57* (17), 4466-4477.
69. Yang, Z. C.; Sun, L. L., Recent technical progress in sample preparation and liquid-phase separation-mass spectrometry for proteomic analysis of mass-limited samples. *Anal. Meth.* **2021,** *13* (10), 1214-1225.
70. Yang, L. W.; George, J.; Wang, J., Deep profiling of cellular heterogeneity by emerging single-cell proteomic technologies. *Proteomics* **2020,** *20* (13), 12.
71. Baxi, A. B.; Pade, L. R.; Nemes, P., Mass spectrometry based proteomics for developmental neurobiology in the amphibian *Xenopus laevis*. In *Amphibian Models of Development and Disease*, Sokol, S. Y., Ed. Elsevier Academic Press Inc: San Diego, 2021; Vol. 145, pp 205-231.
72. Velasquez, E.; Martins-de-Souza, D.; Velasquez, I.; Carneiro, G. R. A.; Schmitt, A.; Falkai, P.; Domont, G. B.; Nogueira, F. C. S., Quantitative subcellular proteomics of the orbitofrontal cortex of schizophrenia patients. *J. Proteome Res.* **2019,** *18* (12), 4240-4253.
73. Hobson, B. D.; Choi, S. J.; Mosharov, E. V.; Soni, R. K.; Sulzer, D.; Sims, P. A., Subcellular proteomics of dopamine neurons in the mouse brain. *eLife* **2022,** *11*.

74. Waas, M.; Kislinger, T., Addressing cellular heterogeneity in cancer through precision proteomics. *J. Proteome Res.* **2020,** *19* (9), 3607-3619.
75. Vo, T. D.; Palsson, B. O., Building the power house: recent advances in mitochondrial studies through proteomics and systems biology. *Am. J. Physiol.-Cell Physiol.* **2007,** *292* (1), C164-C177.
76. Schoof, E. M.; Furtwangler, B.; Uresin, N.; Rapin, N.; Savickas, S.; Gentil, C.; Lechman, E.; Keller, U. A. D.; Dick, J. E.; Porse, B. T., Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **2021,** *12* (1), 15.
77. Fischer, R.; Kessler, B. M., Gel-aided sample preparation (GASP)-A simplified method for gel-assisted proteomic sample generation from protein extracts and intact cells. *Proteomics* **2015,** *15* (7), 1224-1229.
78. HaileMariam, M.; Eguez, R. V.; Singh, H.; Bekele, S.; Ameni, G.; Pieper, R.; Yu, Y. B., S-trap, an ultrafast sample-preparation approach for shotgun proteomics. *J. Proteome Res.* **2018,** *17* (9), 2917-2924.
79. Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M., Universal sample preparation method for proteome analysis. *Nat. Meth.* **2009,** *6* (5), 359-362.
80. Zhang, Z. B.; Dubiak, K. M.; Huber, P. W.; Dovichi, N. J., Miniaturized filter-aided sample preparation (MICRO-FASP) method for high throughput, ultrasensitive proteomics sample preparation reveals proteome asymmetry in *Xenopus laevis* embryos. *Anal. Chem.* **2020,** *92* (7), 5554-5560.
81. 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*, 12.
82. Chen, Q.; Yan, G. Q.; Gao, M. X.; Zhang, X. M., Ultrasensitive proteome profiling for 100 living cells by direct cell injection, online digestion and nano-LC-MS/MS analysis. *Anal. Chem.* **2015,** *87* (13), 6674-6680.
83. Shao, X.; Wang, X. T.; Guan, S.; Lin, H. Z.; Yan, G. Q.; Gao, M. X.; Deng, C. H.; Zhang, X. M., Integrated proteome analysis device for fast single-cell protein profiling. *Anal. Chem.* **2018,** *90* (23), 14003-14010.
84. 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. Ed.* **2018,** *57* (38), 12370-12374.
85. 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*, 1–10.
86. Zhu, Y.; Scheibinger, M.; Ellwanger, D. C.; Krey, J. F.; Choi, D.; Kelly, R. T.; Heller, S.; Barr-Gillespie, P. G., Single-cell proteomics reveals changes in expression during hair-cell development. *eLife* **2019,** *8*, 26.
87. Woo, J.; Williams, S. M.; Markillie, L. M.; Feng, S.; Tsai, C. F.; Aguilera-Vazquez, V.; Sontag, R. L.; Moore, R. J.; Hu, D. H.; Mehta, H. S.; Cantlon-Bruce, J.; Liu, T.; Adkins, J. N.; Smith, R. D.; Clair, G. C.; Pasa-Tolic, L.; Zhu, Y., High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *Nat. Commun.* **2021,** *12* (1), 1.
88. Slavov, N., Scaling up single-cell proteomics. *Mol. Cell. Proteomics* **2022,** *21* (1), 7.

89. Liang, Y. R.; Acor, H.; McCown, M. A.; Nwosu, A. J.; Boekweg, H.; Axtell, N. B.; Truong, T.; Cong, Y. Z.; Payne, S. H.; Kelly, R. T., Fully automated sample processing and analysis workflow for low-input proteome profiling. *Anal. Chem.* **2021,** *93* (3), 1658-1666.
90. Catherman, A. D.; Skinner, O. S.; Kelleher, N. L., Top down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* **2014,** *445* (4), 683-693.
91. Lubeckyj, R. A.; Sun, L. L., Laser capture microdissection-capillary zone electrophoresis-tandem mass spectrometry (LCM-CZE-MS/MS) for spatially resolved top-down proteomics: a pilot study of zebrafish brain. *Mol. Omics* **2022,** *18* (2), 112-122.
92. Smith, L. M.; Agar, J. N.; Chamot-Rooke, J.; Danis, P. O.; Ge, Y.; Loo, J. A.; Pasa-Tolic, L.; Tsybin, Y. O.; Kelleher, N. L.; Consortium Top-Down, P., The human proteoform project: defining the human proteome. *Sci. Adv.* **2021,** *7* (46), 8.
93. 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.
94. Comi, T. J.; Makurath, M. A.; Philip, M. C.; Rubakhin, S. S.; Sweedler, J. V., MALDI MS guided liquid microjunction extraction for capillary electrophoresis-electrospray ionization ms analysis of single pancreatic islet cells. *Anal. Chem.* **2017,** *89* (14), 7765-7772.
95. 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.
96. Kelly, R. T., Single-cell proteomics: progress and prospects. *Mol. Cell. Proteomics* **2020,** *19* (11), 1739-1748.
97. Hook, V.; Lietz, C. B.; Podvin, S.; Cajka, T.; Fiehn, O., Diversity of neuropeptide cell-cell signaling molecules generated by proteolytic processing revealed by neuropeptidomics mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2018,** *29* (5), 807-816.
98. Zhang, L. W.; Khattar, N.; Kemenes, I.; Kemenes, G.; Zrinyi, Z.; Pirger, Z.; Vertes, A., Subcellular peptide localization in single identified neurons by capillary microsampling mass spectrometry. *Sci. Rep.* **2018,** *8*, 10.
99. Reyzer, M. L.; Caprioli, R. M., Imaging mass spectrometry of proteins and peptides. In *Mass Spectrometry for Drug Discovery and Drug Development*, Desiderio, D. M.; Nibbering, N. M. M., Eds. John Wiley & Sons, Inc.: 2013.
100. Chatterji, B.; Pich, A., MALDI imaging mass spectrometry and analysis of endogenous peptides. *Expert Rev. Proteomics* **2013,** *10* (4), 381-388.
101. DeLaney, K.; Phetsanthad, A.; Li, L. J., Advances in high-resolution MALDI mass spectrometry for neurobiology. *Mass Spectrom, Rev.* **2022,** *41* (2), 194-214.
102. Yanes, O.; Tautenhahn, R.; Patti, G. J.; Siuzdak, G., Expanding coverage of the metabolome for global metabolite profiling. *Anal. Chem.* **2011,** *83* (6), 2152-2161.
103. 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.
104. 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.

105. 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.
106. Wu, H. K.; Wheeler, A.; Zare, R. N., Chemical cytometry on a picoliter-scale integrated microfluidic chip. *Proc. Natl Acad. Sci. U. S. A.* **2004,** *101* (35), 12809-12813.
107. Aydogan, C.; Rigano, F.; Krcmova, L. K.; Chung, D. S.; Macka, M.; Mondello, L., Miniaturized LC in molecular omics. *Anal. Chem.* **2020,** *92* (17), 11485-11497.
108. Broeckhoven, K.; Desmet, G., Advances and innovations in liquid chromatography stationary phase supports. *Anal. Chem.* **2021,** *93* (1), 257-272.
109. Mihailova, A.; Malerod, H.; Wilson, S. R.; Karaszewski, B.; Hauser, R.; Lundanes, E.; Greibrokk, T., Improving the resolution of neuropeptides in rat brain with on-line HILIC-RP compared to on-line SCX-RP. *J. Sep. Sci.* **2008,** *31* (3), 459-467.
110. Schirmer, E. C.; Yates 3rd, J. R.; Gerace, L., MudPIT: A powerful proteomics tool for discovery. *Discov. Med.* **2003,** *3* (18), 38-39.
111. Batth, T. S.; Francavilla, C.; Olsen, J. V., Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *J. Proteome Res.* **2014,** *13* (12), 6176-6186.
112. 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.
113. Williams, S. M.; Liyu, A. V.; Tsai, C. F.; Moore, R. J.; Orton, D. J.; Chrisler, W. B.; Gaffrey, M. J.; Liu, T.; Smith, R. D.; Kelly, R. T.; Pasa-Tolic, L.; Zhu, Y., Automated coupling of nanodroplet sample preparation with liquid chromatography-mass spectrometry for high-throughput single-cell proteomics. *Anal. Chem.* **2020,** *92* (15), 10588-10596.
114. Desmet, G.; de Beeck, J. O.; Van Raemdonck, G.; Van Mol, K.; Claerebout, B.; Van Landuyt, N.; Jacobs, P., Separation efficiency kinetics of capillary flow micro-pillar array columns for liquid chromatography. *J. Chromatogr. A* **2020,** *1626*, 11.
115. Yi, L.; Piehowski, P. D.; Shi, T. J.; Smith, R. D.; Qian, W. J., Advances in microscale separations towards nanoproteomics applications. *J. Chromatogr. A* **2017,** *1523*, 40-48.
116. Turiak, L.; Toth, G.; Ozohanics, O.; Revesz, A.; Acs, A.; Vekey, K.; Zaia, J.; Drahos, L., Sensitive method for glycosaminoglycan analysis of tissue sections. *J. Chromatogr. A* **2018,** *1544*, 41-48.
117. Maes, K.; Van Liefferinge, J.; Viaene, J.; Van Schoors, J.; Van Wanseele, Y.; Bechade, G.; Chambers, E. E.; Morren, H.; Michotte, Y.; Heyden, Y. V.; Claereboudt, J.; Smolders, I.; Van Eeckhauta, A., Improved sensitivity of the nano ultra-high performance liquid chromatography-tandem mass spectrometric analysis of low-concentrated neuropeptides by reducing aspecific adsorption and optimizing the injection solvent. *J. Chromatogr. A* **2014,** *1360*, 217-228.
118. Zemenova, J.; Sykora, D.; Freislebenova, A.; Maletinska, L., LC-MS/MS analysis of lipidized analogs of prolactin-releasing peptide utilizing a monolithic column and simple sample preparation. *Bioanalysis* **2017,** *9* (17), 1319-1328.
119. Gregus, M.; Kostas, J. C.; Ray, S.; Abbatiello, S. E.; Ivanov, A. R., Improved sensitivity of ultralow flow LC-MS-based proteomic profiling of limited samples using monolithic capillary columns and FAIMS technology. *Anal. Chem.* **2020,** *92* (21), 14702-14712.

120. Xiang, P. L.; Zhu, Y.; Yang, Y.; Zhao, Z. T.; Williams, S. M.; Moore, R. J.; Kelly, R. T.; Smith, R. D.; Liu, S. R., Picoflow liquid chromatography-mass spectrometry for ultrasensitive bottom-up proteomics using 2-μm-i.d. open tubular columns. *Anal. Chem.* **2020,** *92* (7), 4711-4715.
121. Nemes, P.; Marginean, I.; Vertes, A., Spraying mode effect on droplet formation and ion chemistry in electrosprays. *Anal. Chem.* **2007,** *79* (8), 3105-3116.
122. Bongaerts, J.; Segers, K.; Van Oudenhove, L.; Van Wanseele, Y.; Van Hulle, M.; De Bundel, D.; Mangelings, D.; Smolders, I.; Vander Heyden, Y.; Van Eeckhaut, A., A comparative study of UniSpray and electrospray sources for the ionization of neuropeptides in liquid chromatography tandem mass spectrometry. *J. Chromatogr. A* **2020,** *1628*, 14.
123. Kenderdine, T.; Xia, Z. J.; Williams, E. R.; Fabris, D., Submicrometer nanospray emitters provide new insights into the mechanism of cation adduction to anionic oligonucleotides. *Anal. Chem.* **2018,** *90* (22), 13541-13548.
124. Osbourn, D. M.; Weiss, D. J.; Lunte, C. E., On-line preconcentration methods for capillary electrophoresis. *Electrophoresis* **2000,** *21* (14), 2768-2779.
125. McCool, E. N.; Lubeckyj, R. A.; Shen, X. J.; Chen, D. Y.; Kou, Q.; Liu, X. W.; Sun, L. L., Deep top-down proteomics using capillary zone electrophoresis tandem mass spectrometry: identification of 5700 proteoforms from the *Escherichia coil* proteome. *Anal. Chem.* **2018,** *90* (9), 5529-5533.
126. DeLaney, K.; Jia, D.; 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**, in print.
127. Lindenburg, P. W.; Haselberg, R.; Rozing, G.; Ramautar, R., Developments in interfacing designs for CE-MS: towards enabling tools for proteomics and metabolomics. *Chromatographia* **2015,** *78* (5-6), 367-377.
128. Gomes, F. P.; Yates, III, Recent trends of capillary electrophoresis-mass spectrometry in proteomics research. *Mass Spectrom. Rev.* **2019,** *38* (6), 445-460.
129. Moini, M., Simplifying CE-MS operation. 2. Interfacing low-flow separation techniques to mass spectrometry using a porous tip. *Anal. Chem.* **2007,** *79* (11), 4241-4246.
130. Bonvin, G.; Veuthey, J. L.; Rudaz, S.; Schappler, J., Evaluation of a sheathless nanospray interface based on a porous tip sprayer for CE-ESI-MS coupling. *Electrophoresis* **2012,** *33* (4), 552-562.
131. 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.* **2017,** *28* (4), 597-607.
132. Peuchen, E. H.; Zhu, G. J.; Sun, L. L.; Dovichi, N. J., Evaluation of a commercial electro-kinetically pumped sheath-flow nanospray interface coupled to an automated capillary zone electrophoresis system. *Anal. Bioanal. Chem.* **2017,** *409* (7), 1789-1795.
133. Hirayama, A.; Abe, H.; Yamaguchi, N.; Tabata, S.; Tomita, M.; Soga, T., Development of a sheathless CE-ESI-MS interface. *Electrophoresis* **2018,** *39* (11), 1382-1389.
134. Vermeire, P. J.; Van Schepdael, A.; Petersen, N. J., Development of a novel sheathless CE-ESI-MS interface via a CO_2 laser ablated opening. *Talanta* **2020,** *214*, 8.
135. Xu, X.; Liu, K.; Fan, Z. H., Microscale 2D separation systems for proteomic analysis. *Expert Rev. Proteomics* **2012,** *9* (2), 135-147.

136. Yang, Z. C.; Shen, X. J.; Chen, D. Y.; Sun, L. L., Improved nanoflow RPLC-CZE-MS/MS system with high peak capacity and sensitivity for nanogram bottom-up proteomics. *J. Proteome Res.* **2019,** *18* (11), 4046-4054.
137. Fouque, K. J. D.; Garabedian, A.; Porter, J.; Baird, M.; Pang, X. Q.; Williams, T. D.; Li, L. J.; Shvartsburg, A.; Fernandez-Lima, F., Fast and effective ion mobility-mass spectrometry separation of D-aAmino-acid-containing peptides. *Anal. Chem.* **2017,** *89* (21), 11787-11794.
138. Pfammatter, S.; Bonneil, E.; McManus, F. P.; Prasad, S.; Bailey, D. J.; Belford, M.; Dunyach, J. J.; Thibault, P., A novel differential ion mobility device expands the depth of proteome coverage and the sensitivity of multiplex proteomic measurements. *Mol. Cell. Proteomics* **2018,** *17* (10), 2051-2067.
139. Bailey, D. J.; McDevitt, M. T.; Westphall, M. S.; Pagliarini, D. J.; Coon, J. J., Intelligent data acquisition blends targeted and discovery methods. *J. Proteome Res.* **2014,** *13* (4), 2152-2161.
140. DeLaney, K.; Li, L. J., Data independent acquisition mass spectrometry method for improved neuropeptidomic coverage in crustacean neural tissue extracts. *Anal. Chem.* **2019,** *91* (8), 5150-5158.
141. Lai, X.; Wang, L.; Witzmann, F. A., Issues and applications in label-free quantitative mass spectrometry. *Int. J. Proteomics* **2013,** *2013*, 756039.
142. Dupree, E. J.; Jayathirtha, M.; Yorkey, H.; Mihasan, M.; Petre, B. A.; Dark, C. C., A critical review of bottom-up proteomics: the good, the bad, and the future of this field. *Proteomes* **2020,** *8* (3), 26.
143. Slavov, N., Single-cell protein analysis by mass spectrometry. *Curr. Opin. Chem. Biol.* **2021,** *60*, 1-9.
144. Arul, A. B.; Robinson, R. A. S., Sample multiplexing strategies in quantitative proteomics. *Anal. Chem.* **2019,** *91* (1), 178-189.
145. Lombard-Banek, C.; Choi, S. B.; Nemes, P., Single-cell proteomics in complex tissues using microprobe capillary electrophoresis mass spectrometry. In *Enzyme Activity in Single Cells*, Allbritton, N. L.; Kovarik, M. L., Eds. Academic Press Ltd-Elsevier Science Ltd: London, 2019; Vol. 628, pp 263-292.
146. Marginean, I.; Kelly, R. T.; Page, J. S.; Tang, K. Q.; Smith, R. D., Electrospray characteristic curves: In pursuit of improved performance in the nanoflow regime. *Anal. Chem.* **2007,** *79* (21), 8030-8036.
147. 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.
148. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J., The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Meth.* **2016,** *13* (9), 731-740.
149. Demsar, J.; Curk, T.; Erjavec, A.; Gorup, C.; Hocevar, T.; Milutinovic, M.; Mozina, M.; Polajnar, M.; Toplak, M.; Staric, A.; Stajdohar, M.; Umek, L.; Zagar, L.; Zbontar, J.; Zitnik, M.; Zupan, B., Orange: data mining toolbox in python. *J. Mach. Learn. Res.* **2013,** *14*, 2349-2353.
150. Wu, L. L.; Liu, F.; Cai, H. M., IOAT: an interactive tool for statistical analysis of omics data and clinical data. *BMC Bioinform.* **2021,** *22* (1), 14.

151. Sun, W.; Wu, S. Z.; Wang, X. R.; Zheng, D. X.; Gao, Y. H., An analysis of protein abundance suppression in data dependent liquid chromatography and tandem mass spectrometry with tryptic peptide mixtures of five known proteins. *Eur. J. Mass Spectrom.* **2005,** *11* (6), 575-580.
152. Motoyama, A.; Yates, J. R., Multidimensional LC separations in shotgun proteomics. *Anal. Chem.* **2008,** *80* (19), 7187-7193.
153. Pernemalm, M.; Sandberg, A.; Zhu, Y. F.; Boekel, J.; Tamburro, D.; Schwenk, J. M.; Bjork, A.; Wahren-Herlenius, M.; Amark, H.; Ostenson, C. G.; Westgren, M.; Lehtio, J., In-depth human plasma proteome analysis captures tissue proteins and transfer of protein variants across the placenta. *eLife* **2019,** *8*, No. 24.
154. Carotenuto, R.; Pallotta, M. M.; Tussellino, M.; Fogliano, C., *Xenopus laevis (*Daudin, 1802) as a model organism for bioscience: A historic review and perspective. *Biology-Basel* **2023,** *12* (6), No. 22.
155. Veldman, M. B.; Lin, S., Zebrafish as a developmental model organism for pediatric research. *Pediatr. Res.* **2008,** *64* (5), 470-476.
156. Davey, M. G.; Tickle, C., The chicken as a model for embryonic development. *Cytogenet. Genome Res.* **2007,** *117* (1-4), 231-239.
157. Wiley, H. S.; Wallace, R. A., The structure of vitellogenin. Multiple vitellogenings in *Xenopus laevis* give rise to multiple forms of the yok proteins. *J. Biol. Chem.* **1981,** *256* (16), 8626-8634.
158. Jorgensen, P.; Steen, J. A. J.; Steen, H.; Kirschner, M. W., The mechanism and pattern of yolk consumption provide insight into embryonic nutrition in Xenopus. *Development* **2009,** *136* (9), 1539-1548.
159. 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.
160. 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.
161. Fonslow, B. R.; Stein, B. D.; Webb, K. J.; Xu, T.; Choi, J.; Park, S. K.; Yates, J. R., Digestion and depletion of abundant proteins improves proteomic coverage. *Nat. Methods* **2013,** *10* (1), 54-U129.
162. Fonslow, B. R.; Hixon, M. S.; Stein, B. D.; Webb, K. J.; Xu, T.; Choi, J.; Park, S. K.; Yates, J. R., Addendum: Digestion and depletion of abundant proteins improves proteomic coverage. *Nat. Methods* **2014,** *11* (3), 347–348.
163. Haudenschild, D. R.; Eldridge, A.; Lein, P. J.; Chromy, B. A., High abundant protein removal from rodent blood for biomarker discovery. *Biochem. Biophys. Res. Commun.* **2014,** *455* (1-2), 84-89.
164. Viode, A.; van Zalm, P.; Smolen, K. K.; Fatou, B.; Stevenson, D.; Jha, M.; Levy, O.; Steen, J.; Steen, H.; Network, I., A simple, time- and cost-effective, high-throughput depletion strategy for deep plasma proteomics. *Science Advances* **2023,** *9*, No. eadf9717.
165. Maccarrone, G.; Milfay, D.; Birg, I.; Rosenhagen, M.; Holsboer, F.; Grimm, R.; Bailey, J.; Zolotarjova, N.; Turck, C. W., Mining the human cerebrospinal fluid proteome by immunodepletion and shotgun mass spectrometry. *Electrophoresis* **2004,** *25* (14), 2402-2412.

166. Gong, Y.; Li, X.; Yang, B.; Ying, W. T.; Li, D.; Zhang, Y. J.; Dai, S. J.; Cai, Y.; Wang, J. L.; He, F. C.; Qian, X. H., Different immunoaffinity fractionation strategies to characterize the human plasma proteome. *J. Proteome Res.* **2006,** *5* (6), 1379-1387.
167. Zhu, G. J.; Zhao, P.; Deng, N.; Tao, D. Y.; Sun, L. L.; Liang, Z.; Zhang, L. H.; Zhang, Y. K., Single chain variable fragment displaying M13 phage library functionalized magnetic microsphere based protein equalizer for human serum protein analysis. *Anal. Chem.* **2012,** *84* (18), 7633-7637.
168. Tu, C. J.; Rudnick, P. A.; Martinez, M. Y.; Cheek, K. L.; Stein, S. E.; Slebos, R. J. C.; Liebler, D. C., Depletion of abundant plasma proteins and limitations of plasma proteomics. *J. Proteome Res.* **2010,** *9* (10), 4982-4991.
169. Cao, X. F.; Sandberg, A.; Araujo, J. E.; Cvetkovski, F.; Berglund, E.; Eriksson, L. E.; Pernemalm, M., Evaluation of spin columns for human plasma depletion to facilitate MS-based proteomics analysis of plasma. *J. Proteome Res.* **2021,** *20* (9), 4610-4620.
170. Gao, M. X.; Zhang, J.; Deng, C. H.; Yang, P. Y.; Zhang, X. M., Novel strategy of high-abundance protein depletion using multidimensional liquid chromatography. *J. Proteome Res.* **2006,** *5* (10), 2853-2860.
171. Lundberg, E.; Borner, G. H. H., Spatial proteomics: a powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* **2019,** *20* (5), 285-302.
172. Joshi, R. N.; Stadler, C.; Lehmann, R.; Lehtio, J.; Tegner, J.; Schmidt, A.; Vesterlund, M., TcellSubC: An atlas of the subcellular proteome of human T cells. *Front. Immunol.* **2019,** *10*, No. 2708.
173. McLaughlin, K. L.; Kew, K. A.; McClung, J. M.; Fisher-Wellman, K. H., Subcellular proteomics combined with bioenergetic phenotyping reveals protein biomarkers of respiratory insufficiency in the setting of proofreading-deficient mitochondrial polymerase. *Sci. Rep.* **2020,** *10*, No. 3603.
174. 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), No. eaar5780.
175. Macklin, A.; Khan, S.; Kislinger, T., Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. *Clin. Proteomics* **2020,** *17*, No. 17.
176. Woo, J.; Williams, S. M.; Markillie, L. M.; Feng, S.; Tsai, C. F.; Aguilera-Vazquez, V.; Sontag, R. L.; Moore, R. J.; Hu, D. H.; Mehta, H. S.; Cantlon-Bruce, J.; Liu, T.; Adkins, J. N.; Smith, R. D.; Clair, G. C.; Pasa-Tolic, L.; Zhu, Y., High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *Nat. Commun.* **2021,** *12* (1), No. 6246.
177. Xu, K. R.; Liang, Y. R.; Piehowski, P. D.; Dou, M. W.; Schwarz, K. C.; Zhao, R.; Sontag, R. L.; Moore, R. J.; Zhu, Y.; Kelly, R. T., Benchtop-compatible sample processing workflow for proteome profiling of < 100 mammalian cells. *Anal. Bioanal. Chem.* **2019,** *411* (19), 4587-4596.
178. Weke, K.; Singh, A.; Uwugiaren, N.; Alfaro, J. A.; Wang, T. J.; Hupp, T. R.; O'Neill, J. R.; Vojtesek, B.; Goodlett, D. R.; Williams, S. M.; Zhou, M. W.; Kelly, R. T.; Zhu, Y.; Dapic, I., MicroPOTS analysis of barrett's esophageal cell line models identifies proteomic changes after physiologic and radiation Stress. *J. Proteome Res.* **2021,** *20* (5), 2195-2205.

179. Wisniewski, J. R.; Zougman, A.; Mann, M., Combination of fasp and stagetip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J. Proteome Res.* **2009,** *8* (12), 5674-5678.
180. Budnik, B.; Levy, E.; Harmange, G.; Slavov, N., SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biology* **2018,** *19*, No. 16.
181. 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 transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biology* **2021,** *22* (1), No. 50.
182. Sive, H. L.; Grainger, R. M.; Harland, R. M., *Early development of Xenopus laevis: a laboratory manual*. Cold Spring Harbor Laboratory Press: 2000.
183. Wuhr, M.; Freeman, R. M., Jr.; 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.
184. Nieuwkoop, P.; Faber, J., *Normal table of Xenopus laevis (daudin): a systematical & chronological survey of the development from the fertiilized egg till the end of metamorphosis*. Garland Science: 1994; p 282.
185. 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.
186. Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M. S.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Durkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdzija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L. F.; Wang, M. X.; Oresic, M.; Weng, J. K.; Bocker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X. X.; Pluskal, T., Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **2023,** *41* (4), 447-449.
187. Mi, H. Y.; Ebert, D.; Muruganujan, A.; Mills, C.; Albou, L. P.; Mushayamaha, T.; Thomas, P. D., PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **2021,** *49* (D1), D394-D403.
188. Perez-Riverol, Y.; Csordas, A.; Bai, J. W.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Perez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaino, 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.
189. Pade, L. R.; Stepler, K. E.; Portero, E. P.; DeLaney, K.; Nemes, P., Biological mass spectrometry enables spatiotemporal 'omics: From tissues to cells to organelles. *Mass Spectrom. Rev.* **2023**, No. e21824.
190. Peshkin, L.; Wuhr, M.; Pearl, E.; Haas, W.; Freeman, R. M.; Gerhart, J. C.; Klein, A. M.; Horb, M.; Gygi, S. P.; Kirschner, M. W., On the relationship of protein and mRNA dynamics in vertebrate embryonic development. *Dev. Cell* **2015,** *35* (3), 383-394.
191. Alexander, N. F.; Andrea, M.; Michael, S. L.; Martin, W., Quantitative proteome dynamics across embryogenesis in a model chordate. *bioRxiv* **2023**, 2023.10.04.559613.

192. Slavov, N., Single-cell protein analysis by mass-spectrometry. *Curr. Opin. Chem. Biol.* **2021,** *60*, 19.
193. Liu, D. T.; Yang, S.; Kavdia, K.; Sifford, J. M.; Wu, Z. P.; Xie, B. E.; Wang, Z.; Pagala, V. R.; Wang, H.; Yu, K. W.; Dey, K. K.; High, A. A.; Serrano, G. E.; Beach, T. G.; Peng, J. M., Deep Profiling of Microgram-Scale Proteome by Tandem Mass Tag Mass Spectrometry. *J. Proteome Res.* **2021,** *20* (1), 337-345.
194. Truong, T.; Webber, K. G. I.; Johnston, S. M.; Boekweg, H.; Lindgren, C. M.; Liang, Y. R.; Nydegger, A.; Xie, X. F.; Tsang, T. M.; Jayatunge, D.; Andersen, J. L.; Payne, S. H.; Kelly, R. T., Data-dependent acquisition with precursor coisolation improves proteome coverage and measurement throughput for label-free single-cell proteomics. *Angew. Chem. Int. Ed.* **2023,** *135*, No. e202303415.
195. Kitata, R. B.; Yang, J. C.; Chen, Y. J., Advances in data-independent acquisition mass spectrometry towards comprehensive digital proteome landscape. *Mass Spectrom. Rev.* **2022**, e21781.
196. Pade, L. R.; Lombard-Banek, C.; Li, J.; Nemes, P., Dilute to enrich for deeper proteomics: A yolk-depleted carrier for limited populations of embryonic (frog) cells. *J. Proteome Res.* **2024,** *23* (2), 692–703.
197. 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*.
198. Li, Z. Y.; Huang, M.; Wang, X. K.; Zhu, Y.; Li, J. S.; Wong, C. C. L.; Fang, Q., Nanoliter-scale oil-air-droplet chip-based single cell proteomic analysis. *Anal. Chem.* **2018,** *90* (8), 5430-5438.
199. 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.
200. Tsai, C. F.; Zhang, P. F.; Scholten, D.; Martin, K.; Wang, Y. T.; Zhao, R.; Chrisler, W. B.; Patel, D. B.; Dou, M. W.; Jia, Y. Z.; Reduzzi, C.; Liu, X.; Moore, R. J.; Burnum-Johnson, K. E.; Lin, M. H.; Hsu, C. C.; Jacobs, J. M.; Kagan, J.; Srivastava, S.; Rodland, K. D.; Wiley, H. S.; Qian, W. J.; Smith, R. D.; Zhu, Y.; Cristofanilli, M.; Liu, T.; Liu, H. P.; Shi, T. J., Surfactant-assisted one-pot sample preparation for label-free single-cell proteomics. *Communications Biology* **2021,** *4* (1), 12.
201. Sive, H. L.; Grainger, R. M.; Harland, R. M., *Early development of Xenopus laevis: a laboratory manual*. Cold Spring Harbor Laboratory Press: 2000; p i.
202. Nieuwkoop, P. D.; Faber, J., *Normal table of Xenopus laevis (daudin): a systematical & chronological survey of the development from the fertiilized egg till the end of metamorphosis*. Garland Science: 1994; p 282.
203. Xia, J.; Sinelnikov, I. V.; Han, B.; Wishart, D. S., MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res.* **2015,** *43* (W1), W251-W257.
204. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J., The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **2016,** *13* (9), 731-740.
205. Ma, Y. Y.; Sun, Z. Y.; de Matos, R.; Zhang, J.; Odunsi, K.; Lin, B. Y., Towards an animal model of ovarian cancer: Cataloging chicken blood proteins using combinatorial

peptide ligand libraries coupled with shotgun proteomic analysis for translational research. *Omics-a Journal of Integrative Biology* **2014,** *18* (5), 280-297.
206. Choi, J. O.; Fan, C. D.; Kim, D.; Sharif, M.; An, H.; Park, Y., Elucidating the transactivation domain of the pleiotropic transcription factor Myrf. *Sci. Rep.* **2018,** *8*, 9.
207. Dick, A.; Risau, W.; Drexler, H., Expression of Smad1 and Smad2 during embryogenesis suggests a role in organ development. *Dev. Dyn.* **1998,** *211* (4), 293-305.
208. Hertveldt, V.; Louryan, S.; van Reeth, T.; Dreze, P.; van Vooren, P.; Szpirer, J.; Szpirer, C., The development of several organs and appendages is impaired in mice lacking sp6. *Dev. Dyn.* **2008,** *237* (4), 883-892.
209. Powers, S. E.; Taniguchi, K.; Yen, W. W.; Melhuish, T. A.; Shen, J.; Walsh, C. A.; Sutherland, A. E.; Wotton, D., Tgif1 and Tgif2 regulate Nodal signaling and are required for gastrulation. *Development* **2010,** *137* (2), 249-259.
210. Wu, J.; Li, J.; Chen, K.; Liu, G.; Zhou, Y.; Chen, W.; Zhu, X.; Ni, T. T.; Zhang, B.; Jin, D.; Li, D.; Kang, L.; Wu, Y.; Zhu, P.; Xie, P.; Zhong, T. P., Atf7ip and Setdb1 interaction orchestrates the hematopoietic stem and progenitor cell state with diverse lineage differentiation. *Proceedings of the National Academy of Sciences* **2023,** *120* (1), e2209062120.
211. You, Y.; Cui, Y. D.; Li, Y. T.; Zhang, M.; Wang, X.; Ji, J. X.; Zhang, X. B.; Zhou, M.; Zhang, Z. L.; Ye, S. D.; Wang, X. X., Inhibition of MTA2 and MTA3 induces mesendoderm specification of human embryonic stem cells. *Biochem. Biophys. Res. Commun.* **2021,** *552*, 142-149.
212. Heasman, J., Patterning the early *Xenopus* embryo. *Development* **2006,** *133* (7), 1205-1217.
213. Anderson, C.; Stern, C. D., Chapter Twenty-Six - Organizers in Development. In *Current Topics in Developmental Biology*, Wassarman, P. M., Ed. Academic Press: 2016; Vol. 117, pp 435-454.
214. Arias, A. M.; Steventon, B., On the nature and function of organizers. *Development* **2018,** *145* (5), 10.
215. Kumar, V.; Park, S.; Lee, U.; Kim, J., The Organizer and its signaling in embryonic development. *J. Dev. Biol.* **2021,** *9* (4), No. 47.
216. De Robertis, E. M., Spemann's organizer and the self-regulation of embryonic fields. *Mech. Dev.* **2009,** *126* (11-12), 925-941.
217. and, R. H.; Gerhart, J., Formation and function of spemann's organizer. *Annu. Rev. Cell Dev. Biol.* **1997,** *13* (1), 611-667.
218. Smith, W. C.; Harland, R. M., Expression cloning of Noggin, a new dorsalising factor localized to the Spemann Organizer in *Xenopus* embryos. *Cell* **1992,** *70* (5), 829-840.
219. Surugihalli, C.; Muralidaran, V.; Ryan, C. E.; Patel, K.; Zhao, D.; Sunny, N. E., Branched-chain amino acids alter cellular redox to induce lipid oxidation and reduce de novo lipogenesis in the liver. *Am. J. Physiol.-Endocrinol. Metab.* **2023,** *324* (4), E299-E313.
220. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P., MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal. Chem.* **2014,** *86* (14), 7150-7158.

221. Baxi, A. B.; Li, J.; Quach, V. M.; Pade, L. R.; Moody, S. A.; Nemes, P., Cell lineage-guided mass spectrometry reveals increased energy metabolism and reactive oxygen species in the vertebrate organizer. *Proc. Natl. Acad. Sci. U. S. A.* **2024,** *121* (6), 8.
222. Oey, N. A.; Den Boer, M. E. J.; Wijburg, F. A.; Vekemans, M.; Augé, J.; Steiner, C.; Wanders, R. J. A.; Waterham, H. R.; Ruiter, J. P. N.; Attié-Bitach, T., Long-chain fatty acid oxidation during early human development. *Pediatr. Res.* **2005,** *57* (6), 755-759.
223. Denis, J. A.; Rochon-Beaucourt, C.; Champon, B.; Pietu, G., Global transcriptional profiling of neural and mesenchymal progenitors derived from human embryonic stem cells reveals alternative developmental signaling pathways. *Stem Cells Dev.* **2011,** *20* (8), 1395-1409.
224. Kaltschmidt, C.; Greiner, J. F. W.; Kaltschmidt, B., The transcription factor NF-κb in stem cells and development. *Cells* **2021,** *10* (8), 17.
225. Lüningschrör, P.; Kaltschmidt, B.; Kaltschmidt, C., Knockdown of IKK1/2 promotes differentiation of mouse embryonic stem cells into neuroectoderm at the expense of mesoderm. *Stem Cell Rev. Rep.* **2012,** *8* (4), 1098-1108.
226. Morgan, M. J.; Liu, Z. G., Crosstalk of reactive oxygen species and NF-κB signaling. *Cell Res.* **2011,** *21* (1), 103-115.
227. Heeg-Truesdell, E.; LaBonne, C., Neural induction in *Xenopus* requires inhibition of Wnt-β-catenin signaling. *Dev. Biol.* **2006,** *298* (1), 71-86.
228. Perrotti, D.; Neviani, P., Protein phosphatase 2A: a target for anticancer therapy. *The Lancet Oncology* **2013,** *14* (6), e229-e238.
229. Schuff, M.; Siegel, D.; Bardine, N.; Oswald, F.; Donow, C.; Knöchel, W., *FoxO* genes are dispensable during gastrulation but required for late embryogenesis in *Xenopus laevis*. *Dev. Biol.* **2010,** *337* (2), 259-273.
230. Bowling, S.; Di Gregorio, A.; Sancho, M.; Pozzi, S.; Aarts, M.; Signore, M.; Schneider, M. D.; Barbera, J. P. M.; Gil, J.; Rodríguez, T. A., P53 and mTOR signalling determine fitness selection through cell competition during early mouse embryonic development. *Nat. Commun.* **2018,** *9*, 12.
231. Yu, L. L.; Wu, M.; Zhu, G. Y.; Xu, Y., Emerging roles of the tumor suppressor p53 in metabolism. *Front. Cell. Dev. Biol.* **2022,** *9*, 5.
232. Maejima, Y.; Sadoshima, J., SUMOylation a novel protein quality control modifier in the heart. *Circ. Res.* **2014,** *115* (8), 686-689.
233. Zhang, X. D.; Boyer, L.; Jin, M. J.; Mertens, J.; Kim, Y. S.; Ma, L.; Ma, L.; Hamm, M.; Gage, F. H.; Hunter, T., Metabolic reprogramming during neuronal differentiation from aerobic glycolysis to neuronal oxidative phosphorylation. *eLife* **2016,** *5*, 25.
234. Labuschagne, C. F.; Cheung, E. C.; Blagih, J.; Domart, M. C.; Vousden, K. H., Cell clustering promotes a metabolic switch that supports metastatic colonization. *Cell Metab.* **2019,** *30* (4), 720-+.
235. Song, C. C.; Xu, F. X.; Ren, Z. L.; Zhang, Y. M.; Meng, Y.; Yang, Y. Q.; Lingadahalli, S.; Cheung, E.; Li, G.; Liu, W. W.; Wan, J. B.; Zhao, Y.; Chen, G. K., Elevated exogenous pyruvate potentiates mesodermal differentiation through metabolic modulation and AMPK/mTOR pathway in human embryonic stem cells. *Stem Cell Rep.* **2019,** *13* (2), 338-351.
236. Malate-Aspartate Shuttle.
237. Boell, E. J.; Needham, J.; Rogers, V., Morphogenesis and metabolism studies with the Cartesian diver ultramicromanometer I anaerobic glycolysis of the regions of the

amphibian gastrula. *Proceedings of the Royal Society Series B-Biological Sciences* **1939,** *127* (848), 322-356.
238. Suzuki, Y.; Hayasaka, R.; Hasebe, M.; Ikeda, S.; Soga, T.; Tomita, M.; Hirayama, A.; Kuroda, H., Comparative metabolomics of small molecules specifically expressed in the dorsal or ventral marginal zones in vertebrate gastrula. *Metabolites* **2022,** *12* (6), 11.
239. Angelopoulos, I.; Gakis, G.; Birmpas, K.; Kyrousi, C.; Habeos, E. E.; Kaplani, K.; Lygerou, Z.; Habeos, I.; Taraviras, S., Metabolic regulation of the neural stem cell fate: Unraveling new connections, establishing new concepts. *Front. Neurosci.* **2022,** *16*, 14.
240. Zhou, W. J.; Zhao, T. T.; Du, J. Y.; Ji, G. Y.; Li, X. Y.; Ji, S. F.; Tian, W. Y.; Wang, X.; Hao, A. J., TIGAR promotes neural stem cell differentiation through acetyl-CoA-mediated histone acetylation. *Cell Death Dis.* **2019,** *10*, 13.
241. Sondheimer, N.; Fang, J. K.; Polyak, E.; Falk, M. J.; Avadhani, N. G., Leucine-rich pentatricopeptide-repeat containing protein regulates mitochondrial transcription. *Biochemistry* **2010,** *49* (35), 7467-7473.
242. Chandel, N. S., Mitochondrial complex III: An essential component of universal oxygen sensing machinery? *Respir. Physiol. Neuro.* **2010,** *174* (3), 175-181.
243. Tirichen, H.; Yaigoub, H.; Xu, W. W.; Wu, C. X.; Li, R. S.; Li, Y. F., Mitochondrial reactive oxygen species and their contribution in chronic kidney disease progression through oxidative stress. *Front. Physiol.* **2021,** *12*, 12.
244. Chatterjee, S.; Sil, P. C., ROS-influenced regulatory cross-talk with wnt signaling pathway during perinatal development. *Front. Mol. Biosci.* **2022,** *9*, 17.
245. Lennicke, C.; Rahn, J.; Lichtenfels, R.; Wessjohann, L. A.; Seliger, B., Hydrogen peroxide - production, fate and role in redox signaling of tumor cells. *Cell Commun. Signal.* **2015,** *13*, 19.
246. Garfinkel, A. M.; Mnatsakanyan, N.; Patel, J. H.; Wills, A. E.; Shteyman, A.; Smith, P. J. S.; Alavian, K. N.; Jonas, E. A.; Khokha, M. K., Mitochondrial leak metabolism induces the spemann-mangold organizer via *Hif-1a* in *Xenopus*. *Dev. Cell* **2023,** *58* (22), 2597-+.
247. Jansen, H. J.; Wacker, S. A.; Bardine, N.; Durston, A. J., The role of the Spemann organizer in anterior-posterior patterning of the trunk. *Mech. Dev.* **2007,** *124* (9-10), 668-681.
248. Surugihalli, C.; Porter, T. E.; Chan, A.; Farley, L. S.; Maguire, M.; Zhang, C.; Kattapuram, N.; Muyyarikkandy, M. S.; Liu, H. C.; Sunny, N. E., Hepatic mitochondrial oxidative metabolism and lipogenesis synergistically adapt to mediate healthy embryonic-to-neonatal transition in chicken. *Sci. Rep.* **2019,** *9*, 14.
249. Sunny, N. E.; Bequette, B. J., Glycerol is a major substrate for glucose, glycogen, and nonessential amino acid synthesis in late-term chicken embryos. *J. Anim. Sci.* **2011,** *89* (12), 3945-3953.
250. Simoes-Costa, M.; Bronner, M. E., Establishing neural crest identity: a gene regulatory recipe. *Development* **2015,** *142* (2), 242-257.
251. Kulesa, P. M.; Bailey, C. M.; Kasemeier-Kulesa, J. C.; McLennan, R., Cranial neural crest migration: New rules for an old road. *Dev. Biol.* **2010,** *344* (2), 543-554.
252. Baltzinger, M.; Ori, M.; Pasqualetti, M.; Nardi, I.; Rijli, F. M., *Hoxa2* knockdown in *Xenopus* results in hyoid to mandibular homeosis. *Dev. Dyn.* **2005,** *234* (4), 858-867.
253. Creuzet, S.; Couly, G.; Le Douarin, N. M., Patterning the neural crest derivatives during development of the vertebrate head: insights from avian studies. *J. Anat.* **2005,** *207* (5), 447-459.

254. Rocha, M.; Singh, N.; Ahsan, K.; Beiriger, A.; Prince, V. E., Neural crest development: insights from the zebrafish. *Dev. Dyn.* **2020,** *249* (1), 88-111.
255. Ericsson, R.; Ziermann, J. M.; Piekarski, N.; Schubert, G.; Joss, J.; Olsson, L., Cell fate and timing in the evolution of neural crest and mesoderm development in the head region of amphibians and lungfishes. *Acta Zool.* **2009,** *90*, 264-272.
256. Yoon, J.; Hwang, Y. S.; Lee, M.; Sun, J.; Cho, H. J.; Knapik, L.; Daar, I. O., TBC1d24-ephrinB2 interaction regulates contact inhibition of locomotion in neural crest cell migration. *Nat. Commun.* **2018,** *9*, 14.
257. MBS salts (10X). *Cold Spring Harbor Protocols* **2007,** *2007* (5), pdb.rec10966.
258. Zhu, T. S.; Sun, R.; Zhang, F. F.; Chen, G. B.; Yi, X.; Ruan, G.; Yuan, C. H.; Zhou, S. G.; Guo, T. N., BatchServer: A Web Server for Batch Effect Evaluation, Visualization, and Correction. *J. Proteome Res.* **2021,** *20* (1), 1079-1086.
259. Shi, Y.; Li, J. J.; Chen, C. J.; Gong, M. Z.; Chen, Y.; Liu, Y. X.; Chen, J.; Li, T. Y.; Song, W. H., 5-mehtyltetrahydrofolate rescues alcohol-induced neural crest cell migration abnormalities. *Mol. Brain* **2014,** *7*, 8.
260. Carmona-Fontaine, C.; Theveneau, E.; Tzekou, A.; Tada, M.; Woods, M.; Page, K. M.; Parsons, M.; Lambris, J. D.; Mayor, R., Complement Fragment C3a Controls Mutual Cell Attraction during Collective Cell Migration. *Dev. Cell* **2011,** *21* (6), 1026-1037.
261. Aleksandr, K.; Mansour, A.; Subham, S.; Vincent, K.; Sofia Medina, R.; Hugo, A.; Richard, H.; Leonid, P.; Anne, H. M.-B., From neural border to migratory stage: A comprehensive single cell roadmap of the timing and regulatory logic driving cranial and vagal neural crest emergence. *bioRxiv* **2022**, 2022.03.23.485460.
262. Wang, D. Z.; Wang, L. Y.; Ren, C. L.; Zhang, P.; Wang, M. M.; Zhang, S. Y., High expression of density-regulated re-initiation and release factor drives tumourigenesis and affects clinical outcome. *Oncol. Lett.* **2019,** *17* (1), 141-148.
263. Grant, P. A.; Yan, B.; Johnson, M. A.; Johnson, D. L. E.; Moody, S. A., Novel animal pole-enriched maternal mrnas are preferentially expressed in neural ectoderm. *Dev. Dyn.* **2014,** *243* (3), 478-496.
264. Forman, T. E.; Dennison, B. J. C.; Fantauzzo, K. A., The role of RNA-binding proteins in vertebrate neural crest and craniofacial development. *J. Dev. Biol.* **2021,** *9* (3), 18.
265. Liang, L. N.; Sun, H.; Zhang, W.; Zhang, M. D.; Yang, X.; Kuang, R.; Zheng, H., Meta-Analysis of EMT Datasets Reveals Different Types of EMT. *PLoS One* **2016,** *11* (6), 22.
266. Zhang, Y.; Zhang, H.; Xiao, Z. Q.; Yuan, G. H.; Yang, G. B., *Ipo7* promotes odontoblastic differentiation and inhibits osteoblastic differentiation through regulation of *Runx2* expression and translocation. *Stem Cells* **2022,** *40* (11), 1020-1030.
267. Mi, H. Y.; Ebert, D.; Muruganujan, A.; Mills, C.; Albou, L. P.; Mushayamaha, T.; Thomas, P. D., PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **2021,** *49* (D1), D394-D403.
268. Boot, M. J.; Steegers-Theunissen, R. P. M.; Poelmann, R. E.; Van Iperen, L.; Lindemans, J.; Gittenberger-de Groot, A. C., Folic acid and homocysteine affect neural crest and neuroepithelial cell outgrowth and differentiation in vitro. *Dev. Dyn.* **2003,** *227* (2), 301-308.
269. Leifeld, L.; Fink, K.; Debska, G.; Fielenbach, M.; Schmitz, V.; Sauerbruch, T.; Spengler, U., Anti-apoptotic function of gelsolin in Fas antibody-induced liver failure in vivo. *Am. J. Pathol.* **2006,** *168* (3), 778-785.

270. Suhara, T.; Kim, H. S.; Kirshenbaum, L. A.; Walsh, K., Suppression of Akt signaling induces Fas ligand expression: Involvement of caspase and jun kinase activation in Akt-mediated Fas ligand regulation. *Mol. Cell. Biol.* **2002,** *22* (2), 680-691.
271. Yu, C. S.; Kim, B. S.; Kim, E. H., FAF1 mediates regulated necrosis through PARP1 activation upon oxidative stress leading to dopaminergic neurodegeneration. *Cell Death Differ.* **2016,** *23* (11), 1873-1885.
272. van Venrooy, S.; Fichtner, D.; Kunz, M.; Wedlich, D.; Gradl, D., Cold-inducible RNA binding protein (CIRP), a novel XTcf-3 specific target gene regulates neural development in *Xenopus*. *BMC Dev. Biol.* **2008,** *8*, 14.
273. Nie, S. Y.; Kee, Y.; Bronner-Fraser, M., Caldesmon regulates actin dynamics to influence cranial neural crest migration in *Xenopus*. *Mol. Biol. Cell* **2011,** *22* (18), 3355-3365.
274. Guo, J.; Huan, T., Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography-Mass Spectrometry Based Untargeted Metabolomics. *Anal. Chem.* **2020,** *92* (12), 8072-8080.
275. Kitata, R. B.; Yang, J. C.; Chen, Y. J., Advances in data-independent acquisition mass spectrometry towards comprehensive digital proteome landscape. *Mass Spectrom. Rev.* **2023,** *42* (6), 2324-2348.
276. Derks, J.; Leduc, A.; Wallmann, G.; Huffman, R. G.; Willetts, M.; Khan, S.; Specht, H.; Ralser, M.; Demichev, V.; Slavov, N., Increasing the throughput of sensitive proteomics by plexDIA. *Nat. Biotechnol.* **2023,** *41* (1), 50-+.
277. Kasarla, S. S.; Flocke, V.; Saw, N. M. T.; Fecke, A.; Sickmann, A.; Gunzer, M.; Flögel, U.; Phapale, P., In-vivo tracking of deuterium metabolism in mouse organs using LC-MS/MS. *J. Chromatogr. A* **2024,** *1717*, 464691.
278. Brenig, K.; Grube, L.; Schwarzländer, M.; Köhrer, K.; Stühler, K.; Poschmann, G., The Proteomic Landscape of Cysteine Oxidation That Underpins Retinoic Acid-Induced Neuronal Differentiation. *J. Proteome Res.* **2020,** *19* (5), 1923-1940.
279. Wang, X. F.; Tong, Y. A.; Giorgianni, F.; Beranova-Giorgianni, S.; Penn, J. S.; Jablonski, M. M., Cellular retinol binding protein 1 modulates photoreceptor outer segment folding in the isolated eye. *Developmental Neurobiology* **2010,** *70* (9), 623-635.
280. Baxi, A. B.; Moody, S. A.; Nemes, P. In *NanoLC-MS based discovery proteomic analysis of the frog inner ear*, 68th ASMS Conference on Mass Spectrometry and Allied Topics Online Meeting, June 1-12,2020; Online Meeting, 2020.
281. Lee-Liu, D.; Sun, L. L.; Dovichi, N. J.; Larrain, J., Quantitative proteomics after spinal cord injury (SCI) in a regenerative and a nonregenerative stage in the frog *Xenopus laevis* *Mol. Cell. Proteomics* **2018,** *17* (4), 592-606.
282. Quach, V. M.; Baxi, A. B.; Nemes, P. In *Proteomic characterization of the spemann organizer in Xenopus laevis (frog) embryos*, 67th ASMS Conference on Mass Spectrometry and Allied Topics 2019.
283. Baxi, A. B.; Quach, V. M.; Li, J.; Nemes, P., Proteo-metabolomic study of the spemann's organizer in the vertebrate (frog) embryo. *17th Annual Conference US HUPO* **2021**.
284. Cagnetta, R.; Frese, C. K.; Shigeoka, T.; Krijgsveld, J.; Holt, C. E., Rapid cue-specific remodeling of the nascent axonal proteome. *Neuron* **2018,** *99* (1), 29-46.
285. Schiapparelli, L. M.; Shah, S. H.; Ma, Y. H.; McClatchy, D. B.; Sharma, P.; Li, J. L.; Yates, J. R.; Goldberg, J. L.; Cline, H. T., The retinal ganglion cell transportome identifies proteins transported to axons and presynaptic compartments in the visual system in vivo. *Cell Reports* **2019,** *28* (7), 1935-1947.

286. 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. Ed.* **2016,** *128* (7), 2500-2504.