HHS Public Access

Author manuscript

Curr Top Dev Biol. Author manuscript.

Mass Spectrometry Based Proteomics for Neurodevelopmental Biology in the Amphibian *Xenopus laevis*

Aparna B. Baxi[1,2], **Leena R. Pade**[1], **Peter Nemes**[1,2,*]

[1]Department of Chemistry & Biochemistry, University of Maryland, College Park, College Park, MD, 20742

[2]Department of Anatomy and Cell Biology, The George Washington University, Washington, DC, 20037

Abstract

Xenopus laevis, a prominent vertebrate model in cell and developmental biology, has been instrumental in studying molecular mechanisms of neural development and disease. Recently, high-resolution mass spectrometry (HRMS), a bioanalytical technology, has expanded the molecular toolbox of protein detection and characterization (proteomics). In this chapter, we begin by overviewing the characteristics as well as advantages and challenges of this biological model and technology. We then delve into discussions on their combined use to aid studies on cell differentiation and development of neural tissues. Finally, we reflect on the emerging integration of proteomics and other 'omic technologies to help generate new knowledge, drive and test new hypotheses, and ultimately, advance the understanding of neural development during states of health and disease.

Keywords

Xenopus; mass spectrometry; proteomics; metabolomics; single cell

A. Introduction

Understanding molecular mechanisms of neural development and their impact on the brain is central to advancing health sciences. The CDC reports that 1 in ~2,700 newborns have life-threatening neural tube defects in the US and their treatment and care amounts to $1.5 billion/year (Arth et al., 2017). The need is high and still unmet to expand our basic understanding of the molecular drivers of normal development and the design of next-generation therapeutics to remedy such disorders. Physical sciences, such as bioanalytical chemistry, fill this gap by innovating tools and technologies to enable the study of neural development using model organisms. The South African clawed frog (*Xenopus laevis*) is one such model that has helped important discoveries in neurodevelopmental biology (Borodinsky, 2017), including neural induction, differentiation, regeneration, and human genetic diseases. Readers interested in these topics are referred to the recent *Xenopus*

[*]**Correspondence to:** Peter Nemes, 0107 Chemistry Building, 8051 Regents Dr., College Park, MD 20742, USA; nemes@umd.edu, (Tel.) 301-405-0373.

Community White Papers (Community, 2020) and reviews on these topics (Blum and Ott, 2018; Duncan and Khokha, 2016; Hwang et al., 2019; Pratt and Khakhalin, 2013; Sater and Moody, 2017).

Characterization of all the biomolecules synthesized is essential to understanding neural development and homeostasis; these molecules range from transcripts to proteins, peptides, and metabolites. Technological advances over the last four decades have enabled the detection and quantification of these molecules with scalability in time and space. Today, molecular amplification by polymerase chain reaction (PCR) and next-generation sequencing (NGS) are routinely used to analyze genes and transcripts at trace levels. Likewise, antibodies equipped with signal amplifying "molecular handles" offer sufficient sensitivity to measure select proteins in low abundance. More recently, high-resolution mass spectrometry (HRMS) has advanced studies in developmental biology by enabling the detection of hundreds to thousands of proteins from complex biological samples (Hashimoto et al., 2019; Lombard-Banek et al., 2017). These studies also allow the identification of genes and gene products for testing function. HRMS may be integrated with tools of cell, developmental, and molecular biology including cell fate tracing (e.g., differentiation, cell lineages), transcription/translation-blocking molecules (e.g., morpholinos, shRNA, miRNA), and gene editing tools (e.g., ZFNs, TALENs, CRISPR-Cas9). Discovery analysis of biomolecules followed by hypothesis driven functional studies bridge the biochemical and functional pillars of systems biology.

In this chapter, we focus on the advantages of HRMS based proteomic studies (proteomics) using *Xenopus* to advance neurodevelopmental biology. We then provide vignettes of biological adventures that capitalized on HRMS to answer important questions in neuroscience. Where pertinent, biological insights are highlighted that were made possible due to the integration of *Xenopus* with HRMS proteomics. We also discuss how, in return, *Xenopus* embryos have spurred the development of ultrasensitive instruments and methodologies for proteomic and metabolomic studies in increasing resolution in time and space (e.g., tissues to single cells). Although the focus of this chapter is on neurodevelopmental biology, HRMS and *Xenopus* have provided important insights in a broader context of biology and health research (Federspiel et al., 2019; Peshkin et al., 2015; Presler et al., 2017; Sun et al., 2014; Wuehr et al., 2014). Likewise, HRMS also supported proteomics in other amphibian species. Most recent topics include organ regeneration using the axolotl and questions pertaining to evolutionary biology using the Chinese giant salamander (Demircan et al., 2020; Geng et al., 2019). We apologize to anyone whose work could not be discussed in this chapter due to space constraints.

B. Mass Spectrometry Proteomics for *Xenopus*

Proteomics by HRMS.

The proteome, the full suite of proteins produced in a biological system, can be used as an efficient descriptor of the molecular phenotype. Because proteins are produced downstream of transcription, they are often influenced by intrinsic and extrinsic events to the system. Post-translational modifications (PTMs), such as phosphorylation, further expand the proteomic landscape carrying out biological functions to "proteoforms" (Fig. 1A). As shown

for the early *Xenopus* embryo in Figure 1B, complex dynamics between gene transcription and translation during development challenges the use of mRNA as a general surrogate for quantifying protein expression (Peshkin et al., 2015; Smits et al., 2014). Directly measuring proteins and their PTMs is essential for developing and testing hypotheses for protein-targeted studies.

HRMS is the leading technology for detecting and quantifying proteins and their PTMs. Mass spectrometers enable the unbiased, direct, and specific detection and quantification of large numbers of proteins (Aebersold et al., 2013). HRMS attains unparalleled molecular specificity by "weighing" mass (m/z value) to sub-millidalton (<10 ppm) accuracy and obtaining sequence-specific information on protein/peptide molecules (Aebersold et al., 2013; Walther and Mann, 2010; Zhang et al., 2013). HRMS also enables the identification and precise localization of PTMs on a protein based on mass measurements.f So far ~400 known PTMs (http://www.unimod.org/modifications_list.php) have been explored using HRMS, including more than 200 in biological contexts spanning from signal transduction to cellular localization (Duan and Walther, 2015; Virag et al., 2020). Figure 1C exemplifies the sequencing of a peptide backbone and accurate localization of phosphorylation (see serine residues) based on gas-phase ion fragmentation using high-resolution tandem MS. For example, HRMS-based quantification revealed dynamic changes in protein phosphorylation as *Xenopus* eggs undergo meiotic exit (Presler et al., 2017). Detection and quantification of broad types of PTMs through HRMS has opened new possibilities in the research of protein function, regulation, and turnover.

HRMS aids scientific rigor in biology studies by supplying rich and high-fidelity quantitative data. Table 1 compares the instrumental and technical characteristics of HRMS with antibody-based technologies from classical molecular biology, such as western blots, immunofluorescence, and ELISA. A statistical survey from 2015 found that poorly characterized and non-specific antibodies have taxed the economy of research by over $350 million annually in the US alone (Bradbury and Pluckthun, 2015). Proteomics by HRMS addresses many challenges of protein detection using antibodies. Because detection based on mass (m/z values) does not require functioning probes, HRMS is not affected by batch-to-batch variability encountered in antibodies (Aebersold et al., 2013; Bradbury and Pluckthun, 2015). It follows that HRMS detection is also not challenged by antibody cross-reactivity with other proteins or lack of antibody reactivity among homologous proteins between species. With ready transferability from one lab to another and the availability of commercial bench-top HRMS instruments, HRMS has become an indispensable instrumental tool of protein research (Aebersold et al., 2013; Bradbury and Pluckthun, 2015).

Workflow for HRMS Proteomics.

Protein measurements using HRMS depend on several interconnected steps (Fig. 2A). The three common approaches in proteomics—they are called bottom-up, middle-down, and top-down— have been the focus of recent reviews (Han et al., 2008; Pino et al., 2020a). Figure 2 presents main steps of the most common, bottom-up (shotgun), approach. In this strategy, proteins are digested into typically 8–30 amino acid long peptides using proteolytic

enzymes, such as TPCK-modified trypsin. The resulting peptides are optionally barcoded to enable multiplexed quantification among multiple samples and desalted using commercially available kits (e.g., C18 ZipTips, Pierce; C18 spin-columns, Thermo Fisher). Peptides are then sequenced by HRMS executing tandem MS. Protein-specific (called proteotypic) peptides are mapped against proteins for identification and quantification (Zhang et al., 2013). The bottom-up proteomic strategy is well established, sensitive, and supported by bench-top mass spectrometers available on the market and a variety of commercially and freely available data analysis software packages.

During bottom-up proteomics, peptides are typically separated. Separation decreases molecular complexity in the temporal dimension, helping HRMS perform sequencing by carrying out tandem MS measurements. Ultra/high-performance liquid chromatography (UPLC/HPLC) and capillary electrophoresis (CE) have facilitated the detection of thousands of peptides in *Xenopus* by reducing chemical complexity temporally to tailor to the duty cycle of HRMS. In UPLC, 200 ng–1 μg of peptide (5–10 μg for replicate analysis) is typically separated based on hydrophobicity on a reversed-phase column (typically, a C18 stationary phase), typically identifying ~1,500 – 3,500 proteins with an ~2-hour LCMS separation experiment (called a "run") (Baxi et al., 2018; Peshkin et al., 2015; Sun et al., 2014). For deeper detection of the proteome, molecular complexity may be reduced via longer separations and/or fractionation off/on-line (e.g., strong cation exchange chromatography) followed by reversed phase LC and HRMS analysis of each fraction. Further details about experimental strategies by LC-HRMS are available in the references (Issaq, 2001; Zhang et al., 2013). Alternatively, using CE, peptides may be resolved based on differences in electrophoretic mobility (Fig. 2). With higher separation power (resolution) and sensitivity, we and others have found CE to complement UPLC for protein amounts limited to <100 ng [reviewed in (DeLaney et al., 2019; Fonslow and Yates, 2009; Lombard-Banek et al., 2017). UPLC and CE have uncovered previously unknown gene translational differences between cells of the embryo (see later), reviewed in (Lombard-Banek et al., 2017). Separation approaches using orthogonal principles of separation can be used sequentially to further simplify sample complexity.

Peptides are detected next. The molecules exiting separation by UPLC/CE must be charged, usually by entraining them into a device performing electrospray ionization (ESI). The ionization source is attached to the mass spectrometer, where the ESI-generated ions are analyzed, and detected. Modern-day instruments offer high mass resolving powers for molecular identifications. At present, the average spectral resolution is ~40,000–100,000 full width at half maximum (FWHM) on commercial time-of-flight and 250,000–1,000,000 FWHM on commercial orbitrap ESI mass spectrometers. These instruments are equipped with compartments where peptide ions can be isolated and broken down (called dissociated or fragmented) into structure-specific fragments via tandem MS (aka. MS/MS or MS^2). The resulting fragment ions are analyzed, typically to accurate mass with <5 ppm error.

Tandem HRMS directly sequences the peptide ions and can identify PTMs. Figure 2 shows the tandem mass spectrum of a peptide signal detected at *m/z* 1747.8683. Upon higher-energy collisional dissociation (HCD), this peptide produced characteristic fragment (called *b*- and *y*-type) ions (see insets). These mass and sequence data were matched to the peptide

sequence: AMDQEITVNPQFVQK (from the enzyme Cop9) with a mass accuracy of 10 ppm and a stringent <1% false discovery rate (FDR). To deepen sequence coverage and help identify PTMs and their location on the peptide backbone, additional types of fragmentation techniques are available to HCD. Collision-induced dissociation (CID), electron transfer dissociation (ETD), electron capture dissociation, (ECD), or their combination (e.g., EThCD) have demonstrated complementary performance for sequence determination [reviewed in (Chu et al., 2015)]. HRMS-MS^2 marries high accuracy and precision for identifying peptides (recall Table 1).

The experimentally detected peptide sequences are mapped to the proteome of the organism under study using bioinformatic tools (software). The proteome is typically predicted *in silico* based on the genome and/or experimentally measured mRNA data. The confidence of identifications is enhanced by filtering spectral matches to a stringent <1% FDR, usually against a reverse-sequence decoy database. Common contaminant proteins are also annotated and removed during this step, such as human keratin proteins (e.g., from personnel hair and skin), *E. coli* proteins (e.g., from bacterial cultures processed in the same lab space), and porcine trypsin (used for digesting proteins). A variety of software packages are available to facilitate bioinformatic data analysis. Open-source (free of charge) applications include Skyline (Pino et al., 2020b), MSFragger (Kong et al., 2017), and MaxQuant (Tyanova et al., 2016). Commercial products, such as ProteomeDiscoverer (ThermoFisher Scientific), provide a user-friendly environment by integrating open-source bioinformatic algorithms and gene ontology tools into a single platform. These and other bioinformatic resources studies are reviewed in (Chen et al., 2020). The resulting list of protein identifications may be imported to knowledge bases to evaluate gene ontology, study pathways/networks, and protein interactions, as well as regulatory mechanisms of neural development and developmental diseases that have been annotated for humans. Popular resources include DAVID (Huang et al., 2009), PantherDB (http://pantherdb.org/), Xenbase (http://www.xenbase.org), STRING (Jensen et al., 2009), and Ingenuity Pathway Analysis (QIAGEN, Redwood City, CA).

There are several ways to quantify proteins by HRMS. The concentration of a protein is in stoichiometric relationship with the concentration of its prototypic peptide. As the peptide is quantified based on ion signal abundance by the mass spectrometer, it is possible to quantify using HRMS any protein between samples that yield proteotypic peptide ion signals. To enhance quantitative precision, calculated protein signals may be normalized to the total, mean, or median signal abundance of the proteins. Alternatively, house-keeping genes with comparable expression may be used for normalization. With “label-free quantification” (LFQ), this workflow enables the relative or absolute quantification of proteins (Cox et al., 2014; Lindemann et al., 2017). Further, specialized “mass tags” (TMT, ThermoFisher Scientific; iTRAQ, Sigma-Aldrich)) permit multiplexing quantification among samples at once, enhancing sensitivity, reproducibility, while reducing instrumental analysis time (see later). Chemical (isobaric mass tags) based labeling strategies and stable isotope labeling (such as SILAC) have enabled the detection of newly translated proteins during neural development *Xenopus* (Ong et al., 2002). Applications of mass tag based and chemical labeling based HRMS approaches to answer questions in neurodevelopmental biology are described later in the chapter. As a result, modern HRMS provides accurate and reproducible

quantification of hundreds-to-thousands of different proteins, without requiring the use of functional probes.

Proteomics in *Xenopus* by HRMS.

Xenopus brings many advantages for HRMS-based proteomics. The embryos can be generated year-round, and a female frog lays about 200–500 embryos each time, thus providing a convenient way to scale up workflows. Another advantage of the embryo is size and protein content. The early cleavage-stage embryo offers ~130 μg of protein, sufficient for high-sensitivity HRMS proteomic measurements (Sun et al., 2016). External fertilization and fast development allow timing proteomics experiments to developmental milestones (Baxi et al., 2018; Peshkin et al., 2015; Sun et al., 2014). In *Xenopus,* pigmentation and stereotypical cell divisions permit ready identification and manipulation of cells. Reproducible fate maps are available for embryos at the 16- (Moody, 1987a, b) and 32-cell (Dale and Slack, 1987) stage. We have integrated this information with ultrasensitive HRMS to peer into proteomes of neural, epidermal, or hindgut fated single cells and cell clones in 8–16-cell embryos (Lombard-Banek et al., 2021; Lombard-Banek et al., 2016a). Moreover, undifferentiated cells of the early embryo can be explanted and cultured in the presence of morphogens to induce cell differentiation. Recently, *Xenopus* animal cap explants helped generate a sufficient number of ciliated cells for deep HRMS-based proteomic analysis of RNA-associated proteins required for ciliary beating (Drew et al., 2020).

Challenges in *Xenopus* Proteomics.

Xenopus is not without challenges for HRMS. *Xenopus laevis* has an allotetraploid genome, in which over 56% of genes have a long and a short copy (gene duplication), with minor sequence variations for the same gene. The proteoforms resulting from this genetic background are anticipated to expand the proteome compared to mammalian animal models (Fig. 1A) (Session et al., 2016). The most recently sequenced version of the genome (version 9.2) reports 46,582 protein-coding genes in comparison to ~30,000 genes in the mouse (Waterston et al., 2002). The ~30,000 genes in the human genome have been estimated to generate over a million proteoforms (Jensen, 2004). Consequently, a given *Xenopus laevis* sample can be expected to yield over a million proteoforms. With a diploid genome containing 21,891 protein coding genes (genome version 10), *Xenopus tropicalis* offers a reduced molecular space and an attractive alternative to test functional significance through gene editing (e.g., CRISPR/Cas9) (Nakayama et al., 2014). Larger size and ~5-times more material in *X. laevis* embryos make this model beneficial for cell biological and single-cell analyses (Harland and Grainger, 2011). During the bottom-up proteomic workflow, tryptic digestion cleaves each protein into multiple peptides, thus expanding the complexity of the proteome for analysis (Figs. 2 and 3A).

Such molecular complexity inherent to bottom-up proteomics taxes detection sensitivity and the success of peptide sequencing in HRMS, which in turn can hinder protein identification. In a routine bottom-up proteomics run, we sequence ~25,000 peptides over 100–120 min of separation. Figure 3A presents one such experiment series from our lab, in which ~50,000 precursor peptide ions were analyzed to be able to ascribe ~50% to peptide sequences (<1% FDR) from ~3,500 proteins (<1% FDR). As shown in Figure 3B, separation of a *Xenopus*

sample over 100 minutes resulted in ~1,500 protein IDs in a particular sample from the study series.

Further, as in other vertebrate models, the concentration of proteins spans over a vast dynamic range. Figure 3C ranks proteins by abundance from our measurement of *X. laevis* whole embryos (at gastrulation stages). The HRMS signal intensities reveal a 6–7 log-order dynamic range of concentration in the embryo. Detection of low-abundance proteins, such as transcription factors, is challenged by abundant signals, which saturate the duty cycle of MS/MS measurements. In the cleavage–to–neurula-stage *Xenopus* embryos, the yolk protein vitellogenin (forms Vtga1, Vtga2, Vtgb1, Vtgb2) constitutes up to ~90% of the proteome. These proteins are packaged in yolk platelets, distributed throughout the cell among all the cells in the embryo (Baxi et al., 2018; Peuchen et al., 2016). Abundant yolk proteins interfere with and diminish the detection of lower abundance proteins. As shown in Figure 3D, pelleting of the yolk platelets via soft centrifugation with sucrose before processing the cell lysate for bottom-up proteomics has enhanced detection of the proteome in the *X. laevis* oocyte (Wuehr et al., 2014) and gastrula (Baxi et al., 2018). An ~70% reduction in yolk protein concentration (called “deyolking”), expanded detection from ~700 proteins to ~1,900 proteins, an ~3-fold improvement in protein identification. Many of the proteins detected after deyolking act as transcription factors, players of signaling pathways, and secondary messengers in the dissected neural ectoderm (Baxi et al., 2018). In single *X. laevis* blastomeres, where deyolking cannot be readily performed due to a limitation in sample size, sensitivity was enhanced by freeing up tandem HRMS duty cycle through programmed exclusion of a portion of tryptic peptide ions resulting from vitellogenin (Lombard-Banek et al., 2019).

C. HRMS Proteomics for Neurodevelopmental Biology in *Xenopus*

HRMS-based proteomics has expanded the bioanalytical toolbox of neurodevelopmental biology. This technology has enabled us to identify and quantify hundreds to thousands of proteins with previously unavailable accuracy, precision, and confidence. This metadata can be used to peer into protein/pathway dynamics and to seek out new insights and non-canonical mechanisms driving neural development. The data generated from untargeted experiments in turn has empowered the design of targeted experiments aimed at determining the significance of genes and gene products in specific biological contexts. In what follows, we highlight select studies that illustrate the study of neurodevelopmental questions using *Xenopus* by applying HRMS proteomic approaches in organs, tissues, cells, and single cells (see Table 2).

Organs.

External development and large batch size make *Xenopus* helpful for studying protein expression during neural and sensory organ development. Dynamic changes can be traced by one of many approaches enabling relative and/or absolute protein quantification (see earlier). Designer mass tags (TMTs, iTRAQ, etc.) allow for the relative quantification of proteins among multiple samples (e.g., tissues, biological conditions, replicates). The approach is known to empower comparative analyses with high accuracy and precision (O’Connell et al.,

2018). We have recently used TMT-based relative quantification to assess proteome dynamics during five key steps of inner ear morphogenesis in *X. laevis* (Baxi et al., 2020). Sustained enrichment of integrin, collagen, and extracellular matrix proteins marked formation of the inner ear.

Tissues.

Tissues may be readily manipulated and dissected from embryos, larvae, and tadpoles for HRMS proteomics. As shown in Figure 4, we recently dissected the newly established neural ectoderm at the onset of gastrulation by tracing the lineage of the neural fated cell in the early embryo (Baxi et al., 2018). This survey of the neural ectodermal proteome captured proteins with critical roles during neural development, including molecular signaling via the FGF, Wnt, RhoGTPase, and Notch pathways. A similarly tissue-specific approach also allowed us to characterize the proteome of the developing Spemann organizer (Quach et al., 2019). This signaling center is transiently established at the onset of gastrulation and plays critical roles in patterning the embryo, including the induction of the neural ectoderm. We uncovered that the Spemann organizer tissues were enriched in mitochondrial proteins that participate in oxidative phosphorylation (Baxi et al., 2021). These and other large-scale proteomic studies from transiently formed tissues generate an important resource of thousands of proteins that can be functionally evaluated for their role in neural induction and development.

Indeed, HRMS and *Xenopus laevis* enabled large-scale proteome dynamics during regeneration after spinal cord injury (Fig. 5). Multiplexed relative proteomic analysis following spinal cord injury in regenerative (early tadpole) and non-regenerative (post-metamorphosis) stages of spinal cords (Fig. 5A) facilitated the identification of 172 proteins that were significantly upregulated during spinal cord regeneration in the early *Xenopus* tadpole (Lee-Liu et al., 2018). Proteomic studies of organogenesis and regeneration from tadpole and later stages of *Xenopus* benefit from an overall decreased abundance of yolk proteins that are consumed during development, contrary to yolk-laden tissues from earlier stages of development studied in the next section.

Cells.

Xenopus tissues can be readily explanted to monitor axon growth along with pulsed metabolic labeling to study protein translation. This labeling approach called pulsed stable isotope labeling by amino acids in cell culture (pSILAC) relies on the integration of heavy isotope labeled amino acids to monitor protein translation. Figure 5B shows one such study in which retinal ganglion cells were explanted in media containing heavy isotope labeled essential amino acids to quantify protein remodeling during the growth of retinal axons in the presence of axon guiding cues (Cagnetta et al., 2018). This study design uncovered nascent proteomes axons give rise to specifically in response to chemo-attractive cues, such as Netrin, BDNF, and Sema3A (Fig. 5B). Results from this study agreed with immunofluorescence imaging that demonstrated that proteins synthesized in response to repulsive gradients (Sema3A) are localized at the tip of the axon growth cone, revealing their direct involvement in axon growth.

Alternatively, bio-orthogonal non-canonical amino acid tagging (BONCAT) enables the quantification of protein synthesis by incorporation of non-canonical amino acids such as L-azidohomoalanine (AHA) in newly synthesized proteins provides. The AHA can be equipped with biotin tags for enrichment and detection by HRMS. AHA was injected into the visual system to label neuronal proteins synthesized in response to different visual cues to study neural plasticity in *Xenopus* tadpoles (Schiapparelli et al., 2019). Ultimately, a group of candidate plasticity proteins were identified that changed significantly between different visual experiences and primarily involved functions such as RNA splicing, protein translation, as well as chromatin remodeling. Teasing apart the nascent proteome through metabolic labeling and untargeted proteomic analysis via HRMS has the potential to yield valuable insights into neural development.

HRMS has also been valuable in confirming and identifying proteins in *Xenopus* (Fig. 6). This technology has sufficient sensitivity and throughput to identify proteins in large numbers of spots from 2D gel electrophoresis. As shown in Figure 6A, a study of photosensory cells in the retina using 2D gel electrophoresis implicated a 16 kDa protein in the organization of the cell membrane. Using HRMS (Fig. 6B), it was possible to make a confident identification for the *Xenopus laevis* retinal protein (CRBP1) based on the high-precision sequencing of proteotypic peptides generated in a bottom-up workflow. Validation of protein identity by HRMS, in turn, permitted follow-up biological experiments to spatially interrogate expression using immunohistochemistry. Targeted knockdown of CRBP1, shown in Figure 6C, revealed the molecular role of organizing the membrane of the cells (Wang et al., 2010). Integration of HRMS with high-resolution imaging and functional experiments can be a viable tool for exploring molecular mechanisms of tissue differentiation in space and time.

Single Cells.

Xenopus has also been used to explore molecular mechanisms with cellular and subcellular resolution (Fig. 7). Large cell sizes in the early developing embryo facilitate cell handling, sample collection, and detection sensitivity. For example, an average cell in the 16-cell *X. laevis* embryo measures ~250 μm diameter (nonspherical cell shape, see Fig. 7A) and contains ~16 μg of protein. With ~90% protein content dominated by yolk proteins, a single cell is estimated to yield ~1 μg of non-yolk proteins. These protein amounts are sufficient for high-sensitivity instruments, opening the door to single-cell proteomics in the developing embryo (Lombard-Banek et al., 2016a; Sun et al., 2016). Further, cells of the early embryo can be identified based on fate maps (see earlier) to target proteomic measurements to specific cell lineages and tissues (Onjiko et al., 2017c). Cell clones may be marked by fluorescent dyes or expressing fluorescent proteins (Fig. 4). The quantitative proteomic studies from identified embryonic cells have revealed previously unknown proteomic cell heterogeneity at a surprisingly early stage of development. This information is useful for follow-up studies to decipher the functional roles of the proteins during early induction of the neural fate.

Results from such experiments can drive new technologies and our current understanding of developmental processes forward. By analyzing the molecular composition of identified

cells in the early stage embryo, we discovered small molecules (called metabolites) that alter normal cell fate decisions between neural and epidermal tissues in the embryo (Onjiko et al., 2015). Using our home-built proteomic HRMS instruments, we have also characterized the proteome state of early embryonic cells that are fated to form neural, ectodermal, and endodermal tissues (Lombard-Banek et al., 2019; Lombard-Banek et al., 2016a; Lombard-Banek et al., 2016b). Figure 7A compares the proteome profiles of the cells (Lombard-Banek et al., 2016a). These data revealed that the chordate embryo sets up proteomic cell asymmetry along the dorsal-ventral and animal-vegetal axes at a rather early stage of development, when cell heterogeneity is only detectable along the animal-vegetal axes by deep sequencing of the transcriptome. Further, the single cells appeared to execute proteomic programs that were indicative of their tissue fate. For example, the neural-destined dorsal cell was enriched in geminin, a protein important for regulating transcription of neuronal-fate promoting genes. These results highlight the enrichment of potentially fate-promoting proteins and metabolites in the neural-fated domain of the embryo early on.

Fast molecular events can be studied using subcellular HRMS. We have recently developed a capillary microsampling approached to swiftly (< 5 s) collect cellular proteins from cells in the live embryo (Lombard-Banek et al., 2021; Lombard-Banek et al., 2019; Onjiko et al., 2017a; Onjiko et al., 2017b, c; Portero and Nemes, 2019). As illustrated in Figure 7B, a small volume (|10 nL) was aspirated from identified cells and analyzed by HRMS, resulting in the identification of ~800 protein groups (representing method sensitivity of |700 zmol). Spatially and temporally scalable proteomic analysis also enabled us to monitor protein reorganization during the formation of the neural tissue fated cell clone (Lombard-Banek et al., 2019). Proteins were group based on their expression profile over the cell clone, revealing progressive and transient mechanisms of protein regulation. These quantitative HRMS datasets provide previously unavailable information on the proteomic state of spatially and temporally evolving clones, supplying candidate genes or gene products to determine function.

Further, subcellular HRMS has also enabled *in vivo* molecular system biology of neural differentiation in live embryos developing to behaving tadpoles (Fig. 7C). We have recently tailored capillary microsampling to minimize damage to the cell and its membrane as a miniscule portion of the cell content was aspirated. The approach caused no damage to the neighboring cell or embryonic development. *In vivo* microsampling with CE-ESI-MS permitted the detection of both proteins and metabolites in single ventral and dorsal cells of the 8-cell embryo (Lombard-Banek et al., 2021). As shown in Figure 7C, the results revealed proteo-metabolomic networks with differential quantitative activities between the neural (D1) and epidermal (V1) fated embryonic cells. The microsampling approach preserved the viability of the sampled cell and the embryo. In our study, 95% of the sampled embryos successfully developed into sentient tadpoles that had indistinguishable anatomy from their wild-type siblings. Further, the tadpoles also presented indistinguishable visual behavior in a background color preference assay that was validated based on double-axotomy of the optic nerves (positive Control, see Fig. 7C). *In vivo* subcellular proteo-metabolomic CE-HRMS with expands the molecular toolbox of molecular systems biology, cell and developmental biology, and neurobiology.

D. Conclusions

Xenopus and HRMS-based proteomics empower developmental neurobiology. This biological model and bioanalytical technology offer important advantages in experimentation and data fidelity. Detection and quantification of large numbers, hundreds-to-thousands, of proteins (and their PTMs) over a broad linear dynamic range of concentration complements molecular information available from transcriptomics to investigate complex molecular dynamics in the developing system. HRMS also enables the detection of small molecules for holistic systems molecular biology, including molecules partaking in energy balance (e.g., adenosine triphosphate), red-ox control (e.g., glutathione and oxidized glutathione), and signaling (e.g., retinoic acid). Additionally, the technology affords an imaging modality (not discussed here) to document the distribution of molecules in variable spatial resolution, as reviewed in (Lombard-Banek et al., 2017). In one such form of HRMS, MALDI-TOF may be used to map neuropeptides in the crab central nervous system (DeLaney et al., 2021) and small molecules in sections of *Xenopus* embryos and tadpoles (Wang et al., 2019). The integration of *Xenopus* with multi-'omics, including HRMS-based proteomics and metabolomics, raises a potential to study molecular mechanisms of development in previously unavailable details at the realms of the organism, organ, tissue, and the cell.

Acknowledgments

This research was partially supported by the National Science Foundation under award IOS-1832968 CAREER (to P.N.) and the National Institutes of Health under awards 1R35GM124755 (to P.N.) and 1F31DC018742 (to A.B.B).

REFERENCES

Aebersold R, Burlingame AL, Bradshaw RA, 2013. Western blots versus selected reaction monitoring assays: time to turn the tables? Mol. Cell. Proteomics 12, 2381–2382.

Arth AC, Tinker SC, Simeone RM, Ailes EC, Cragan JD, Grosse SD, 2017. Inpatient hospitalization costs associated with birth defects among persons of all ages — united states, 2013. MMWR Morb Mortal Wkly Rep 66, 41–46.

Baxi AB, Lombard-Banek C, Moody SA, Nemes P, 2018. Proteomic characterization of the neural ectoderm fated cell clones in the *Xenopus laevis* embryo by high-resolution mass spectrometry. ACS Chemical Neuroscience, 2064–2073.

Baxi AB, Moody SA, Nemes P, 2020. NanoLC-MS based discovery proteomic analysis of the frog inner ear, 68th ASMS Conference on Mass Spectrometry and Allied Topics Online Meeting.

Baxi AB, Quach VM, Li J, Nemes P, 2021. Proteo-metabolomic study of the spemann's organizer in the vertebrate (frog) embryo. 17th Annual Conference US HUPO.

Blum M, Ott T, 2018. *Xenopus*: an undervalued model organism to study and model human genetic disease. Cells Tissues Organs 205, 303–313.

Borodinsky LN, 2017. *Xenopus laevis* as a model organism for the study of spinal cord formation, development, function and regeneration. Front. Neural Circuits 11, 1–9.

Bradbury A, Pluckthun A, 2015. Standardize antibodies used in research. Nature 518, 27–29.

Cagnetta R, Frese CK, Shigeoka T, Krijgsveld J, Holt CE, 2018. Rapid cue-specific remodeling of the nascent axonal proteome. Neuron 99, 29–46.

Chen C, Hou J, Tanner JJ, Cheng JL, 2020. Bioinformatics methods for mass spectrometry-based proteomics data analysis. Int. J. Mol. Sci 21, 25.

Chu IK, Siu CK, Lau JKC, Tang WK, Mu XY, Lai CK, Guo XH, Wang X, Li N, Xia Y, Kong XL, Oh HB, Ryzhov V, Turecek F, Hopkinson AC, Siu KWM, 2015. Proposed nomenclature for peptide ion fragmentation. Int. J. Mass Spectrom 390, 24–27.

Community X, 2020. Xenopus Community White Paper.

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M, 2014. Accurate proteome-wide label-free quantification delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics 13, 2513–2526.

Dale L, Slack JMW, 1987. Fate map for the 32-cell stage of *Xenopus laevis*. Development 99, 527–551.

DeLaney K, Hu M, Hellenbrand T, Dickinson PS,MP,N,L,L, 2021. Mass spectrometry quantification, localization, and discovery of feeding-related neuropeptides in cancer borealis. ACS Chemical Neuroscience Feb 1, Epub ahead of print.

DeLaney K, Sauer CS, Vu NQ, Li LJ, 2019. Recent advances and new perspectives in capillary electrophoresis-mass spectrometry for single cell “omics”. Molecules 24, 1–21.

Demircan T, Sibai M, Altuntas E, 2020. Proteome data to explore the axolotl limb regeneration capacity at neotenic and metamorphic stages. Data Brief 29, 1–7.

Drew K, Lee C, Cox RM, Dang V, Devitt CC, McWhite CD, Papoulas O, Huizar RL, Marcotte EM, Wallingford JB, 2020. A systematic, label-free method for identifying RNA-associated proteins in vivo provides insights into vertebrate ciliary beating machinery. Dev. Biol 467, 108–117.

Duan GY, Walther D, 2015. The roles of post-translational modifications in the context of protein interaction networks. PLoS Comput. Biol 11, 1–23.

Duncan AR, Khokha MK, 2016. *Xenopus* as a model organism for birth defects – congenital heart disease and heterotaxy. Semin. Cell Dev. Biol 51, 1–16.

Federspiel JD, Tandon P, Wilczewski CM, Wasson L, Herring LE, Venkatesh SS, Cristea IM, Conlon FL, 2019. Conservation and divergence of protein pathways in the vertebrate heart. Plos Biology 17, 1–31.

Fonslow BR, Yates JR, 2009. Capillary electrophoresis applied to proteomic analysis. J. Sep. Sci 32, 1175–1188.

Geng XF, Guo JL, Zang XY, Chang CF, Shang HT, Wei H, Xu CS, 2019. Proteomic analysis of eleven tissues in the chinese giant salamander (*Andrias davidianus*). Sci Rep 9, 1–10.

Han XM, Aslanian A, Yates JR, 2008. Mass spectrometry for proteomics. Curr. Opin. Chem. Biol 12, 483–490.

Harland RM, Grainger RM, 2011. Xenopus research: metamorphosed by genetics and genomics. Trends Genet. 27, 507–515.

Hashimoto Y, Greco TM, Cristea IM, 2019. Contribution of mass spectrometry-based proteomics to discoveries in developmental biology, in: Woods AG, Darie CC (Eds.), Advancements of Mass Spectrometry in Biomedical Research, 2nd Edition. Springer International Publishing Ag, Cham, pp. 143–154.

Huang DW, Sherman BT, Lempicki RA, 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research 37, 1–13.

Hwang WY, Marquez J, Khokha MK, 2019. *Xenopus*: driving the discovery of novel genes in patient disease and their underlying pathological mechanisms relevant for organogenesis. Front. Physiol 10, 1–7.

Issaq HJ, 2001. The role of separation science in proteomics research. Electrophoresis 22, 3629–3638.

Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, von Mering C, 2009. STRING 8 — a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Research 37, D412–D416.

Jensen ON, 2004. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Curr. Opin. Chem. Biol 8, 33–41.

Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI, 2017. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. Nat. Methods 14, 513-+.

Lee-Liu D, Sun LL, Dovichi NJ, Larrain J, 2018. Quantitative proteomics after spinal cord injury (SCI) in a regenerative and a nonregenerative stage in the frog *Xenopus laevis* Mol. Cell. Proteomics 17, 592–606.

Lindemann C, Thomanek N, Hundt F, Lerari T, Meyer HE, Wolters D, Marcus K, 2017. Strategies in relative and absolute quantitative mass spectrometry based proteomics. Biol. Chem 398, 687–699.

Lombard-Banek C, Li J, Portero EP, Onjiko RM, Singer CD, Plotnick D, Al Shabeeb RQ, Nemes P, 2021. In vivo subcellular mass spectrometry enables proteo-metabolomic single-cell systems biology in a chordate embryo developing to a normally behaving tadpole (*X. laevis*). Angewandte Chemie-International Edition.

Lombard-Banek C, Moody SA, Manzin MC, Nemes P, 2019. Microsampling capillary electrophoresis mass spectrometry enables single-cell proteomics in complex tissues: developing cell clones in live *Xenopus laevis* and zebrafish embryos. Anal. Chem 91, 4797–4805.

Lombard-Banek C, Moody SA, Nemes P, 2016a. Single-cell mass spectrometry for discovery proteomics: quantifying translational cell heterogeneity in the 16-cell frog (Xenopus) embryo. Angewandte Chemie-International Edition 55, 2454–2458.

Lombard-Banek C, Portero EP, Onjiko RM, Nemes P, 2017. New-generation mass spectrometry expands the toolbox of cell and developmental biology. Genesis 55, 1–34.

Lombard-Banek C, Reddy S, Moody SA, Nemes P, 2016b. 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 15, 2756–2768.

Moody SA, 1987a. Fates of the blastomeres of the 16-cell stage *Xenopus* embryo. Dev. Biol 119, 560–578.

Moody SA, 1987b. Fates of the blastomeres of the 32-cell-stage Xenopus embryo. Dev. Biol 122, 300–319.

Nakayama T, Blitz IL, Fish MB, Odeleye AO, Manohar S, Cho KWY, Grainger RM, 2014. Cas9-based genome editing in *Xenopus tropicalis*, in: Doudna JA, Sontheimer EJ (Eds.), Use of Crispr/Cas9, Zfns, and Talens in Generating Site-Specific Genome Alterations. Elsevier Academic Press Inc, San Diego, pp. 355–375.

O'Connell JD, Paulo JA, O'Brien JJ, Gygi SP, 2018. Proteome-wide evaluation of two common protein quantification methods. J. Proteome Res 17, 1934–1942.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M, 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol. Cell. Proteomics 1, 376–386.

Onjiko RM, Moody SA, Nemes P, 2015. Single-cell mass spectrometry reveals small molecules that affect cell fates in the 16-cell embryo. Proceedings of the National Academy of Sciences of the United States of America 112, 6545–6550.

Onjiko RM, Plotnick DO, Moody SA, Nemes P, 2017a. Metabolic comparison of dorsal versus ventral cells directly in the live 8-cell frog embryo by microprobe single-cell CE-ESI-MS. Analytical Methods 9, 4964–4970.

Onjiko RM, Portero EP, Moody SA, Nemes P, 2017b. In situ microprobe single-cell capillary electrophoresis mass spectrometry: metabolic reorganization in single differentiating cells in the live vertebrate (*Xenopus laevis*) embryo. Anal. Chem 89, 7069–7076.

Onjiko RM, Portero EP, Moody SA, Nemes P, 2017c. Microprobe capillary electrophoresis mass spectrometry for single-cell metabolomics in live frog (*Xenopus laevis*) embryos. J. Vis. Exp, 1–9.

Peshkin L, Wuhr M, Pearl E, Haas W, Freeman RM, Gerhart JC, Klein AM, Horb M, Gygi SP, Kirschner MW, 2015. On the relationship of protein and mRNA dynamics in vertebrate embryonic development. Developmental Cell 35, 383–394.

Peuchen EH, Sun LL, Dovichi NJ, 2016. Optimization and comparison of bottom-up proteomic sample preparation for early-stage *Xenopus laevis* embryos. Anal. Bioanal. Chem 408, 4743–4749.

Pino LK, Rose J, O'Broin A, Shah S, Schilling B, 2020a. Emerging mass spectrometry-based proteomics methodologies for novel biomedical applications. Biochem. Soc. Trans 48, 1953–1966.

Pino LK, Searle BC, Bollinger JG, Nunn B, MacLean B, MacCoss MJ, 2020b. The skyline ecosystem: informatics for quantitative mass spectrometry proteomics. Mass Spectrometry Reviews 39, 229–244.

Portero EP, Nemes P, 2019. Dual cationic-anionic profiling of metabolites in a single identified cell in a live *Xenopus laevis* embryo by microprobe CE-ESI-MS. Analyst 144, 892–899.

Pratt KG, Khakhalin AS, 2013. Modeling human neurodevelopmental disorders in the *Xenopus* tadpole: from mechanisms to therapeutic targets. Disease Models & Mechanisms 6, 1057–1065.

Presler M, Van Itallie E, Klein AM, Kunz R, Coughlin ML, Peshkin L, Gygi SP, Wuhr M, Kirschner MW, 2017. Proteomics of phosphorylation and protein dynamics during fertilization and meiotic exit in the *Xenopus* egg. Proceedings of the National Academy of Sciences of the United States of America 114, E10838–E10847.

Quach VM, Baxi AB, Nemes P, 2019. Proteomic characterization of the spemann organizer in *Xenopus laevis* (frog) embryos, 67th ASMS Conference on Mass Spectrometry and Allied Topics

Sater AK, Moody SA, 2017. Using *Xenopus* to understand human diseases and developmental disorders. Genesis 55, 1–14.

Schiapparelli LM, Shah SH, Ma YH, McClatchy DB, Sharma P, Li JL, Yates JR, Goldberg JL, Cline HT, 2019. The retinal ganglion cell transportome identifies proteins transported to axons and presynaptic compartments in the visual system in vivo. Cell Reports 28, 1935–1947.

Session AM, Uno Y, Kwon T, Hapman JAC, Toyoda A, Takahashi S, Fukui A, Hikosaka A, Suzuki A, Kondo M, van Heeringen SJ, Quigley I, Heinz S, Ogino H, Ochi H, Hellsten U, Lyons JB, Simakov O, Putnam N, Stites J, Kuroki Y, Tanaka T, Michiue T, Watanabe M, Ogdanovic OB, Lister R, Georgiou G, Paranjpe SS, Van Kruijsbergen I, Shu SQ, Carlson J, Kinoshita T, Ohta Y, Mawaribuchi S, Jenkins J, Grimwood J, Schmutz J, Mitros T, Mozaffari SV, Suzuki Y, Haramoto Y, Yamamoto TS, Takagi C, Heald R, Miller K, Haudenschild C, Kitzman J, Nakayama T, Zutsu YI, Robert J, Fortriede J, Burns K, Lotay V, Karimi K, Yasuoka Y, Dichmann DS, Flajnik MF, Houston DW, Shendure J, DuPasquier L, Vize PD, Zorn AM, Ito M, Marcotte EM, Wallingford JB, Ito Y, Asashima M, Ueno N, Matsuda Y, Veenstra GJC, Fujiyama A, Harland RM, Taira M, Rokhsar DS, 2016. Genome evolution in the allotetraploid frog *Xenopus laevis*. Nature 538, 336–343.

Smits AH, Lindeboom RGH, Perino M, van Heeringen SJ, Veenstra GJC, Vermeulen M, 2014. Global absolute quantification reveals tight regulation of protein expression in single *Xenopus* eggs. Nucleic Acids Research 42, 9880–9891.

Sun LL, Bertke MM, Champion MM, Zhu GJ, Huber PW, Dovichi NJ, 2014. Quantitative proteomics of *Xenopus laevis* embryos: expression kinetics of nearly 4000 proteins during early development. Sci Rep 4, 1–9.

Sun LL, Dubiak KM, Peuchen EH, Zhang ZB, Zhu GJ, Huber PW, Dovichi NJ, 2016. Single cell proteomics using frog (*Xenopus laevis*) blastomeres isolated from early stage embryos, which form a geometric progression in protein content. Anal. Chem 88, 6653–6657.

Tyanova S, Temu T, Cox J, 2016. The maxquant computational platform for mass spectrometry-based shotgun proteomics. Nature Protocols 11, 2301–2319.

Virag D, Dalmadi-Kiss B, Vekey K, Drahos L, Klebovich I, Antal I, Ludanyi K, 2020. Current trends in the analysis of post-translational modifications. Chromatographia 83, 1–10.

Walther TC, Mann M, 2010. Mass spectrometry-based proteomics in cell biology. J. Cell Biol 190, 491–500.

Wang M, Dubiak K, Zhang ZB, Huber PW, Chen DDY, Dovichi NJ, 2019. MALDI-imaging of early stage *Xenopus laevis* embryos. Talanta 204, 138–144.

Wang XF, Tong YA, Giorgianni F, Beranova-Giorgianni S, Penn JS, Jablonski MM, 2010. Cellular retinol binding protein 1 modulates photoreceptor outer segment folding in the isolated eye. Developmental Neurobiology 70, 623–635.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J,

Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Larty EO, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Strange-Thomann NS, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Trevaskis E, Tromp J, Ucla C, Vidal AU, Vinson JP, von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES, Mouse Genome Sequencing C, 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562.

Wuehr M, Freeman RM, Presler M, Horb ME, Peshkin L, Gygi SP, Kirschner MW, 2014. Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. Current Biology 24, 1467–1475.

Zhang YY, Fonslow BR, Shan B, Baek MC, Yates JR, 2013. Protein analysis by shotgun/bottom-up proteomics. Chem. Rev 113, 2343–2394.

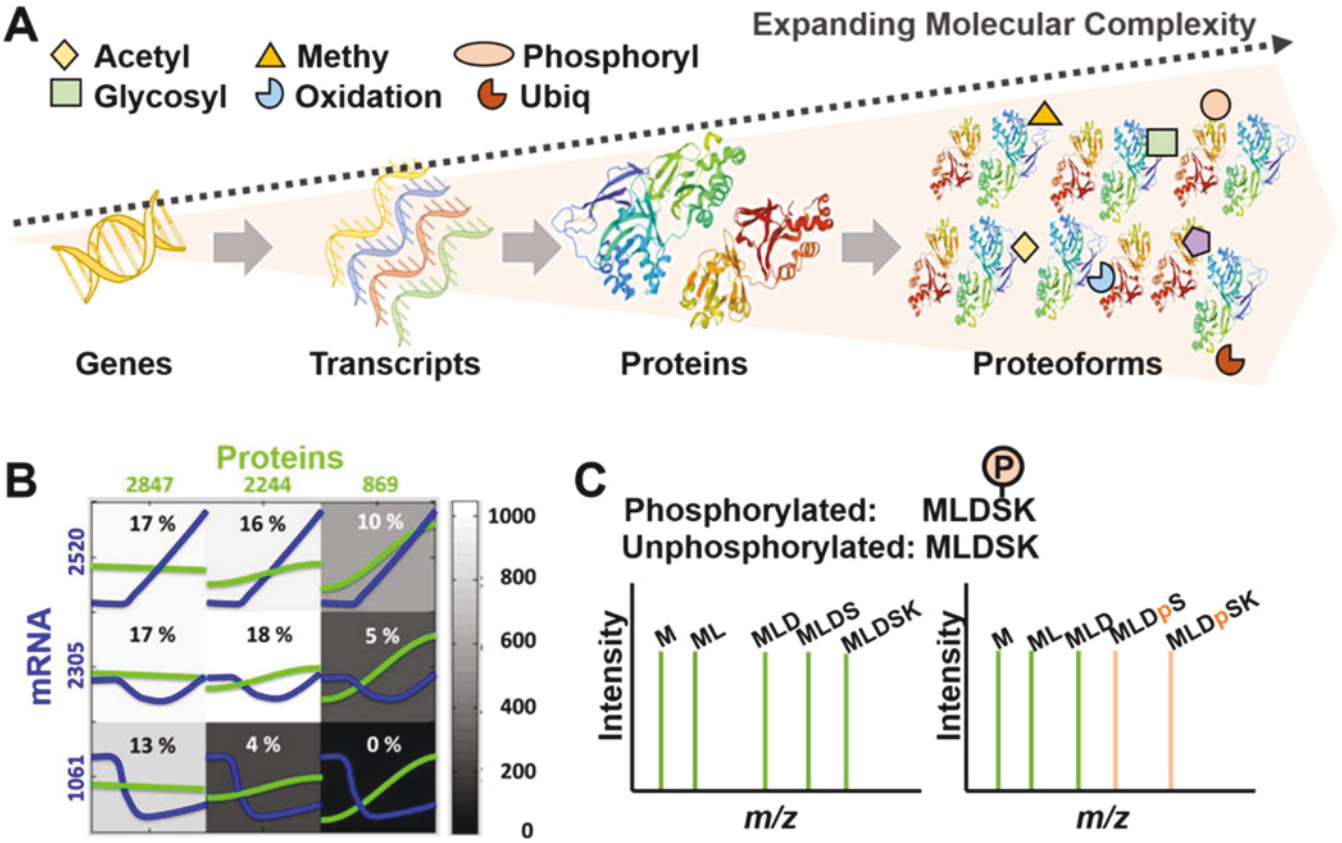

Figure 1.

Measurement of proteins in biological systems. **(A)** Complex proteomic landscape from molecular systems biology, expanded by post-translational modifications (PTMs). Representative PTMs are illustrated. **(B) C**omplex temporal correlation between mRNA and protein abundances in developing systems call for direct measurement of proteins. The example clusters mRNA and protein expression in the *X. laevis* embryo. Number of genes (gray scale, % genes shown), transcripts (blue), and proteins (green) are marked for each cluster. Each cluster marks the number of transcripts, proteins, and the % of gene products belonging. **(C)** High-resolution mass spectrometry (HRMS) provides unmatched specificity and selectivity to identify and quantify proteins, including their PTMs. The example shows the sequencing of a peptide and localization of phosphorylation. Images were adapted with permission from References (Han et al., 2008; Peshkin et al., 2015).

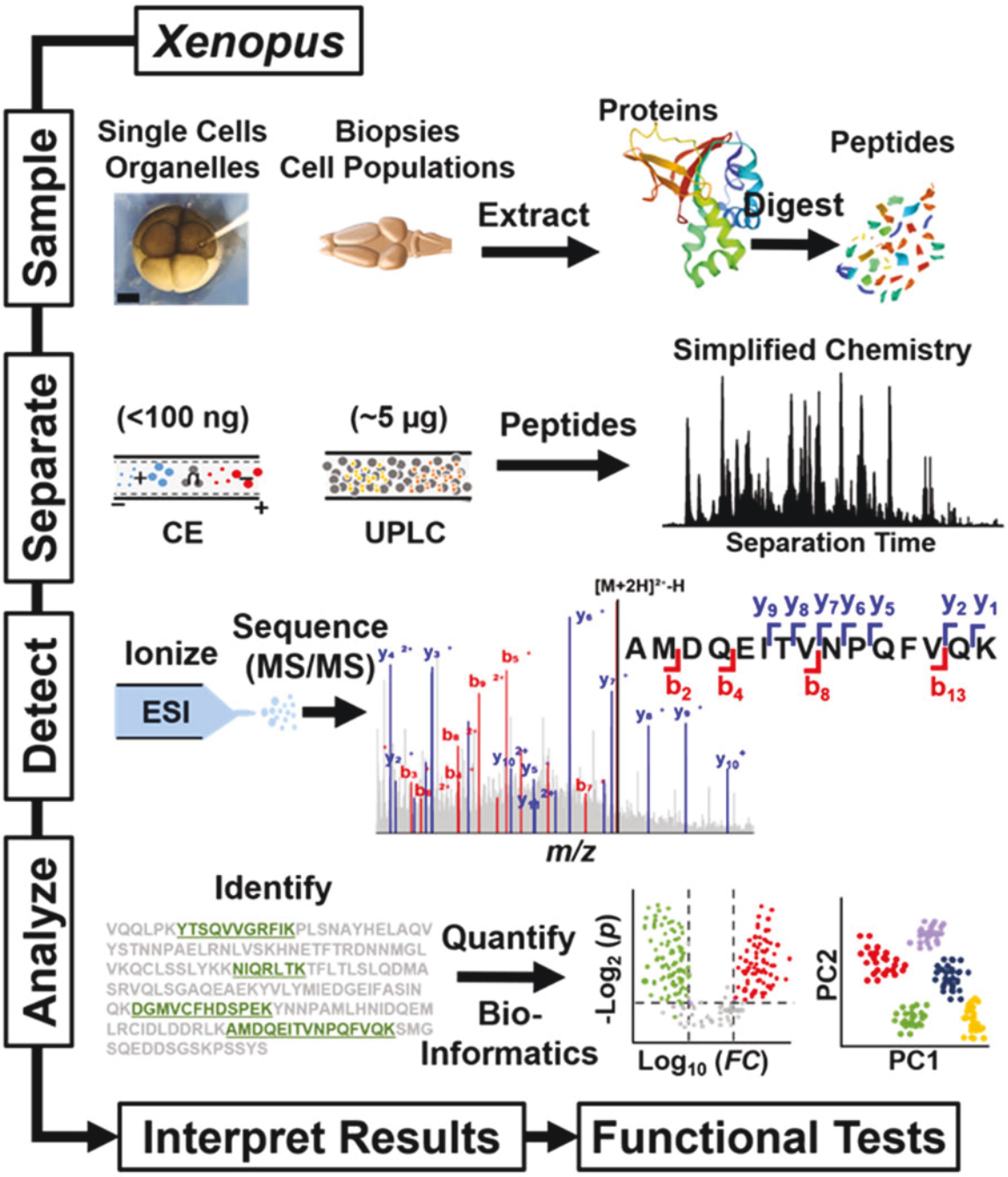

Figure 2.

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 (Onjiko et al., 2017b).

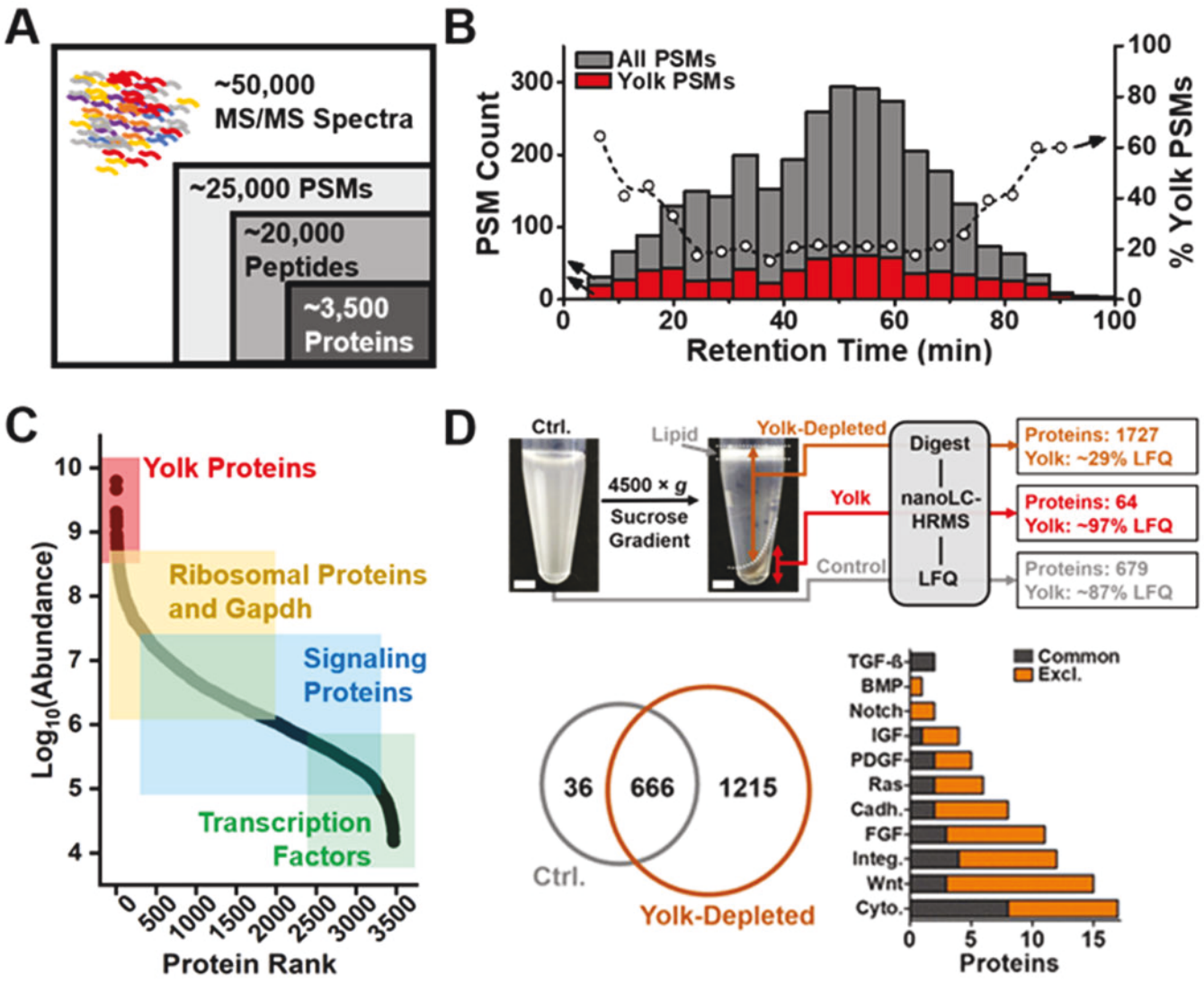

Figure 3.

Challenges and possible solutions in MS proteomics in *Xenopus*. **(A)** Vast molecular complexity in proteomic digests tax the sequencing of peptide signals via tandem MS (MS/MS). **(B)** Chemical separation efficiently reduces molecular complexity to enhance peptide identifications. Abundant yolk peptides limit sequencing throughput. **(C)** A broad dynamic range of concentration challenges the identification of low-abundance proteins. For example, abundant yolk proteins overwhelm signal abundance, hindering detection sensitivity. Shaded boxes represent the types of proteins that are typically present at a given abundance range. **(D)** Removal of yolk platelets **(top panel)** by sucrose pelleting **(bottom panel)** ca. triples the number of identifiable proteins, permitting the quantification of important signaling pathways. Figures were adapted with permission from (Baxi et al., 2018).

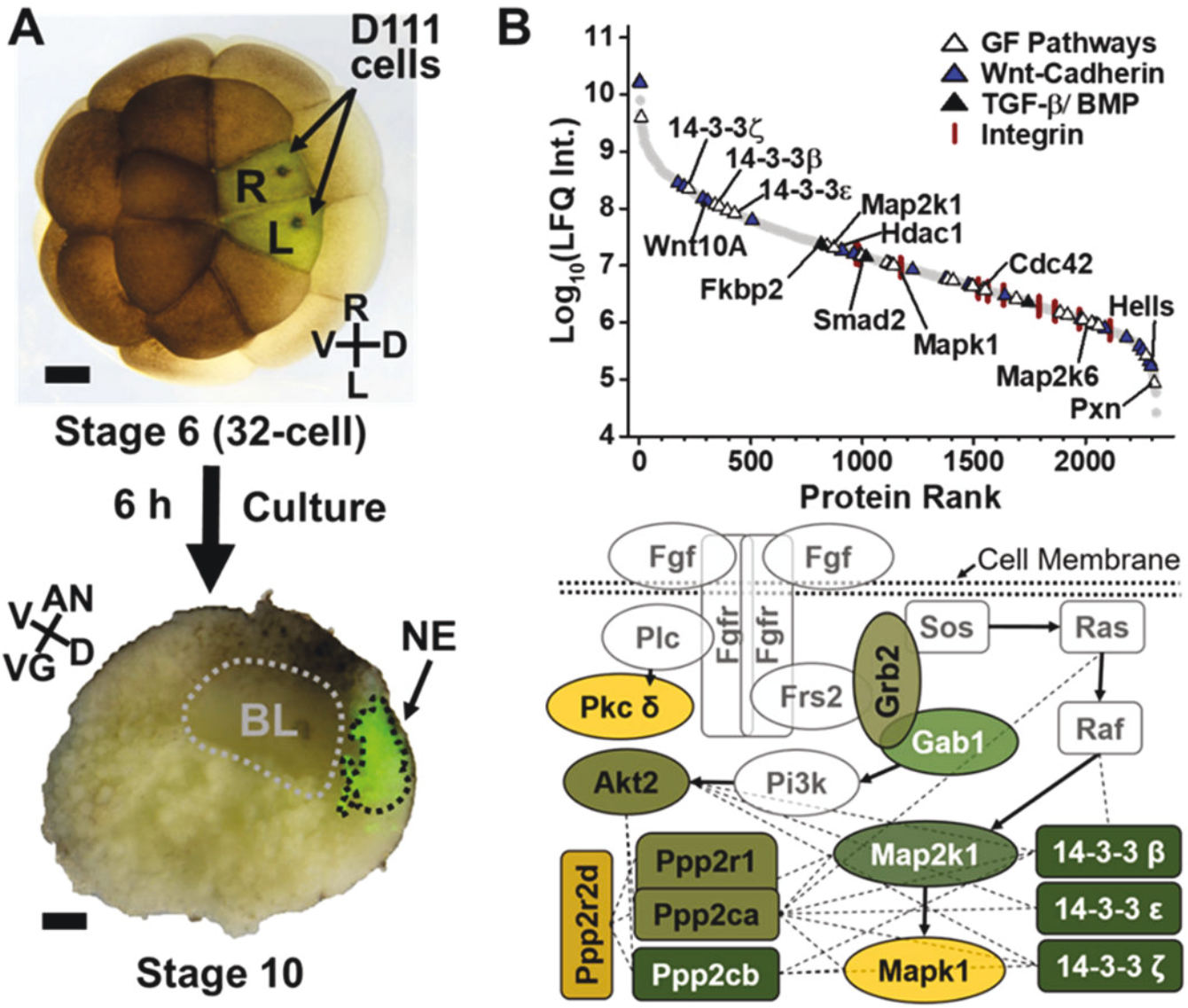

Figure 4.

Lineage tracing and HRMS to profile protein expression during neural induction. **(A)** Injection of neural fated (D111) cells of the *Xenopus* embryo and subsequent dissection of the neural ectoderm at the beginning of gastrulation (stage 10). **(B)** Dynamic range of quantified proteins in the tissue **(top panel)**. Representative proteins with important roles in neurodevelopmental signaling pathways are highlighted. Kinases and second messengers from the FGF pathway are represented **(bottom panel)**. Figures were adapted with permission from (Baxi et al., 2018).

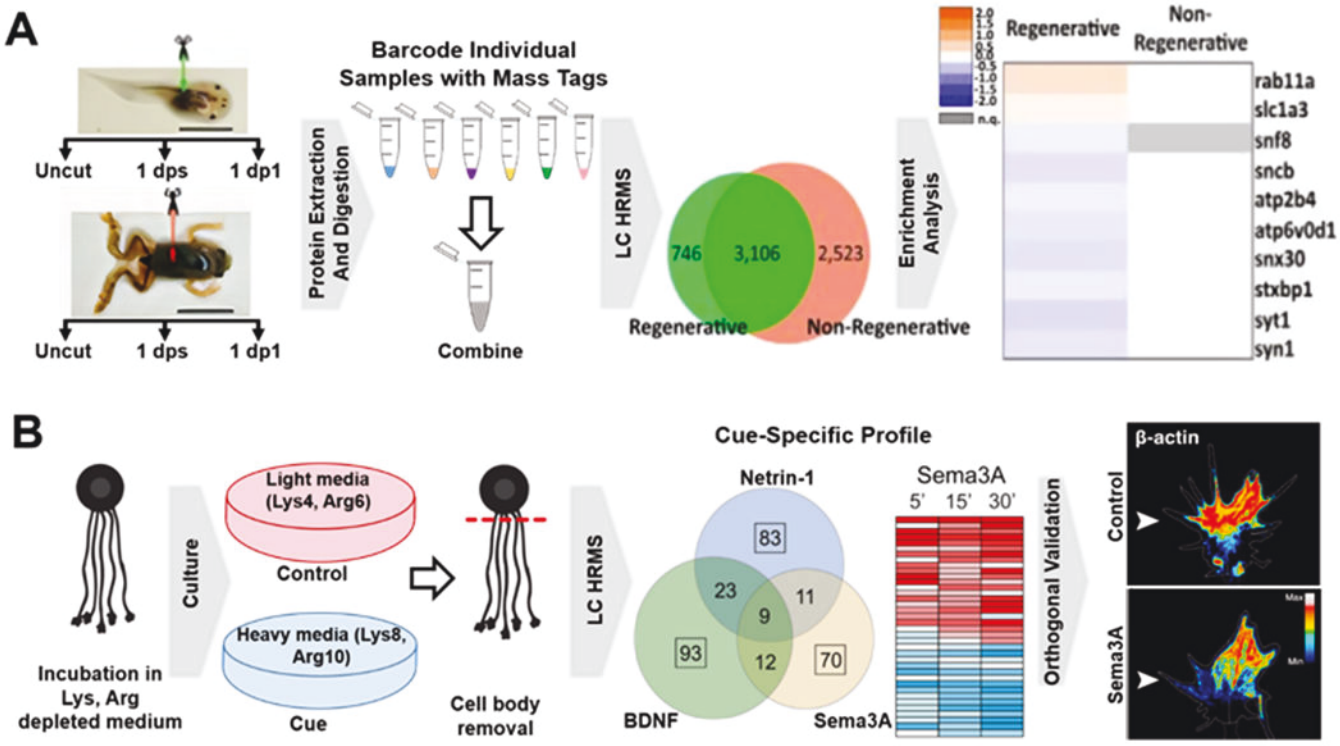

Figure 5.

Relative protein quantification using mass tags and metabolic labeling. **(A)** HRMS was used to compare proteins in regenerative and non-regenerative spinal cord tissues via multiplexing quantification. **(B)** Example of a pulsed SILAC experiment in a neural cell culture to identify proteins synthesized during cue-driven axon guidance. The repulsive gradient with Sema3A was found to lead to an enrichment of beta-actin on the opposite side of the axon growth cone. Figures were adapted with permission from (Cagnetta et al., 2018; Lee-Liu et al., 2018).

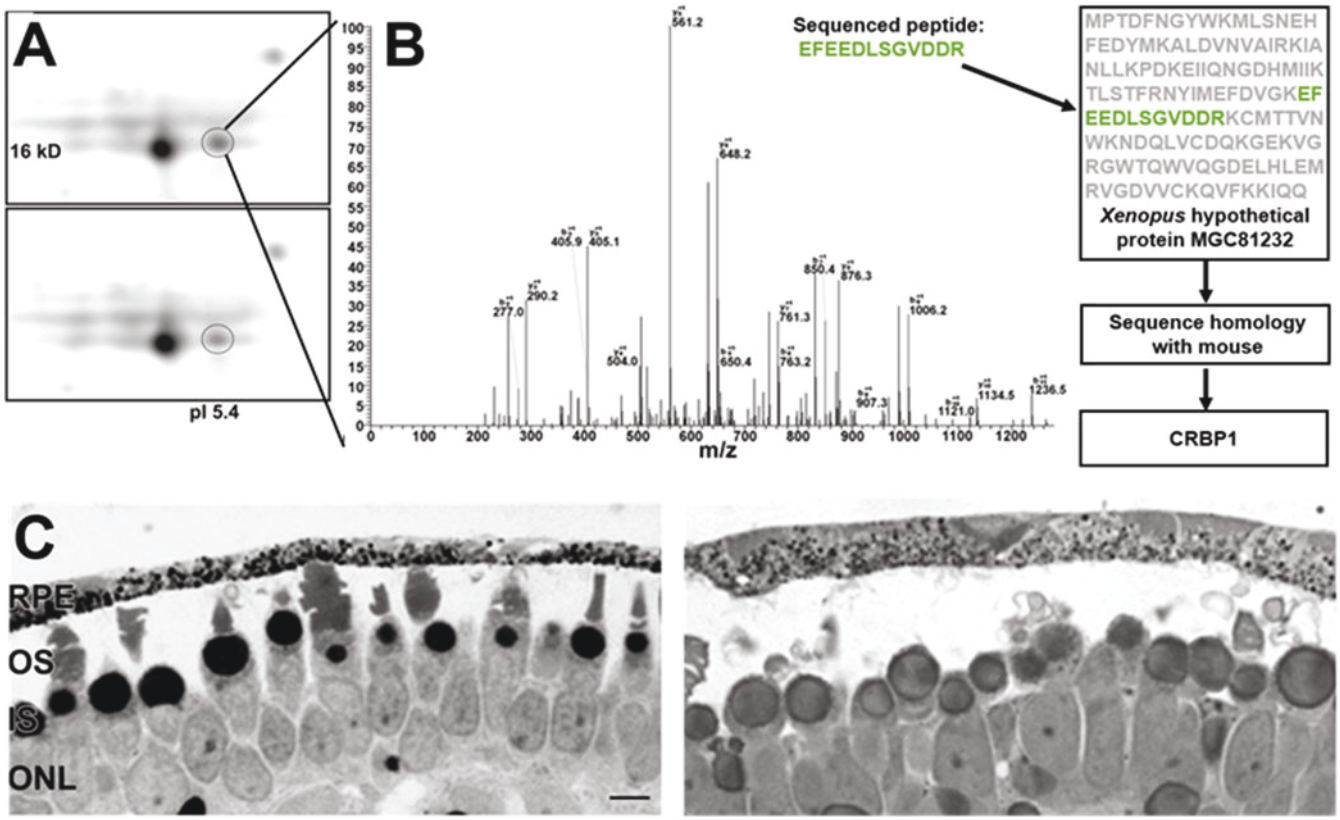

Figure 6.

HRMS enables identification of a previously uncharacterized retinal protein. **(A)** 2D-gel electrophoresis of retinal cells revealing diffuse protein spots, challenging identification. **(B)** MS analysis of the encircled spot identified the *Xenopus laevis* retinal protein (CRBP1) based on sequence homology with the mouse proteome. **(C)** Targeted knock-down of CRBP1 in retina with morpholino **(bottom panel)** compared to control **(top)** show significant disruption of the outer segment. Figures were adapted with permission from (Wang et al., 2010).

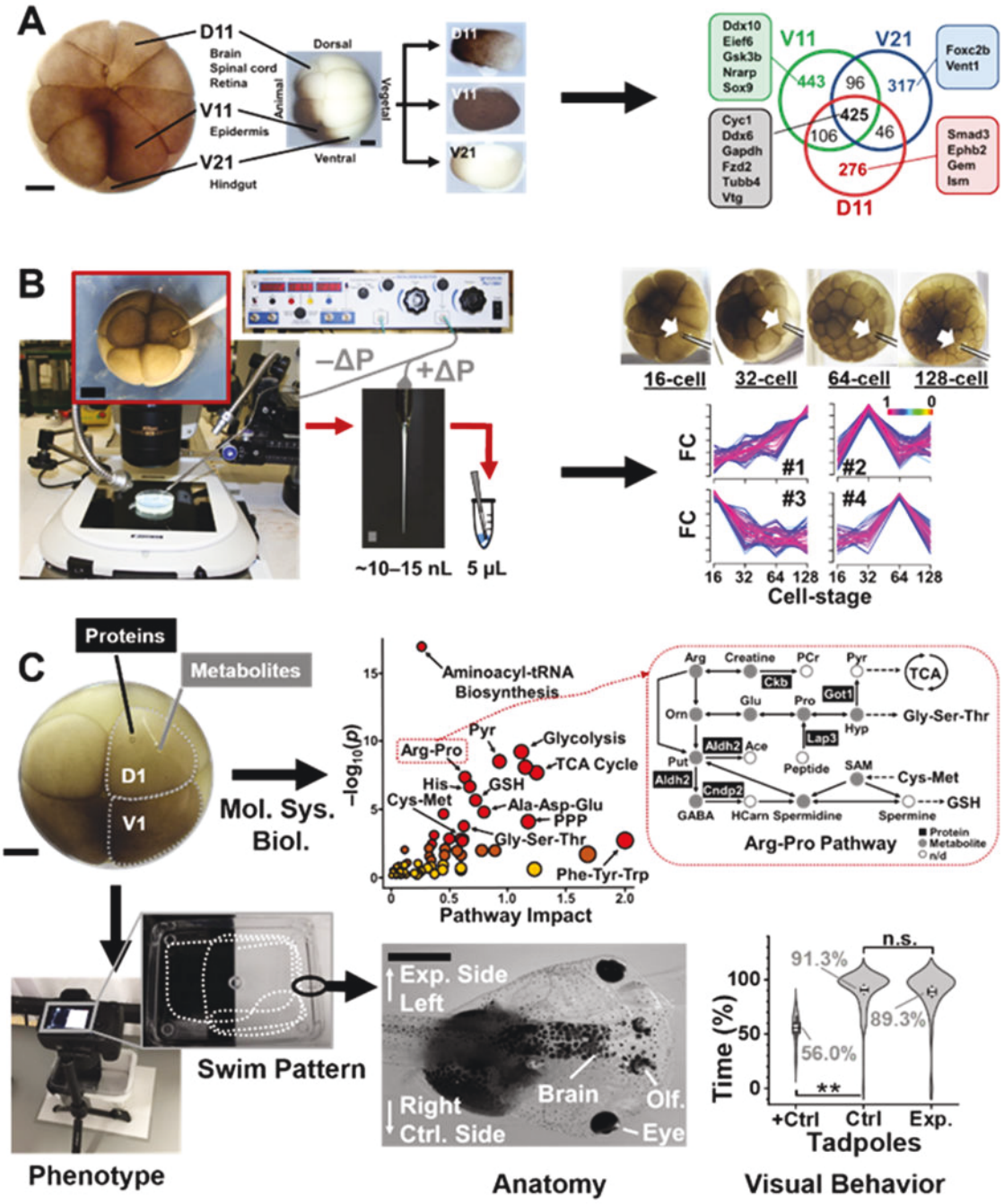

Figure 7.

Single-cell proteo-metabolomics of neural tissue fated cells in live embryos. **(A)** Whole-cell dissection HRMS **(left panel)** revealing proteomic differences between D11, V11, and V21 blastomeres fated to respectively form neural, epidermal, and endodermal tissues **(right panel)**. Scale = 250 µm. **(B)** *In situ* microsampling HRMS of the neural-fated cell **(left panel)** revealing proteome reorganization in descendent cell clones **(right panel)**. Scale = 500 µm (black), 200 µm (gray). **(C)** Microsampling and HRMS detecting proteo-metabolomic networks **(top panel)** in the D1 and V1 cells of live embryos, without affecting

the anatomy or visual behavior of the future tadpoles **(bottom panel)**. Positive control for visual assay (+Ctrl) via double-axotomy of the optic nerves. Scale = 250 μm (embryo), 1 mm (tadpole). Figures adapted with permission from (Lombard-Banek et al., 2021; Lombard-Banek et al., 2019; Lombard-Banek et al., 2016a; Onjiko et al., 2015; Onjiko et al., 2017b).

Table 1.

Comparison of protein detection by high-resolution mass spectrometry (HRMS) and classical antibody-based assays (e.g., Western blot, immunofluorescence, ELISA).

Figure of Merit	HRMS	Antibody-Based Assays
Experimental Strategy	Discovery or targeted	Targeted
Mechanism of Molecular Identification	MS/MS, (accurate) m/z, separation	Antibody detects epitope and (optionally) drives signal amplification reaction
Specificity, Accuracy	Exquisite specificity and accuracy (<5 ppm accuracy routine)	Depends on antibody specificity, requires validation by HRMS
Sensitivity	Fmol–zmol. Detection sensitive to quantitative interferences.	Up to low fmol levels
Quantification	Yes; absolute and relative quantification with precision and accuracy	Semi-quantitative based on a calibration curve (absolute)
Throughput	High throughput: thousands of proteins identified and quantified in a single experiment in ~3 h. Sample multiplexing possible.	Low throughput: usually 1 protein at a time. Limited to no multiplexing.
Spatial information	Yes; MS can be operated as a “mass microscope”	Yes, antibody-based fluorescence imaging
Cost	Affordable via MS Facilities or high start-up cost; no requirement to validate antibodies	Low start-up cost; high consumables cost; requirement to validate antibodies

Table 2.

Examples of proteomics by HRMS supporting neurodevelopmental studies in *Xenopus*.

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	(Wang et al., 2010)
Otic vesicles	High-pH RPLC-Low-pH RPLC-MS	Relative (TMT)	~5,000	(Baxi et al., 2020)
Spinal cord tissues	SCX-RPLC-MS	Relative (iTRAQ)	~6,000	(Lee-Liu et al., 2018)
Neural ectoderm tissues (beginning of gastrulation)	High-pH RPLC-Low-pH RPLC-MS	Label-free	~2,500	(Baxi et al., 2018)
Spemann organizer tissue (beginning of gastrulation)	High-pH RPLC-Low-pH RPLC-MS	Relative (TMT)	~3,000	(Baxi et al., 2021; Quach et al., 2019)
Axons	RPLC-MS	Metabolic labeling (pSILAC)	~350 labeled proteins	(Cagnetta et al., 2018)
Axons	MudPIT-MS	Metabolic labeling (BONCAT)	~4,800	(Schiapparelli et al., 2019)
Neural, epidermal, and endodermal fated single cells	CE-MS	Label-free	~438	(Lombard-Banek et al., 2016a; Lombard-Banek et al., 2016b)
Neural fated single cells, visual assay of the tadpole	CE-MS	Label-free	~738	(Lombard-Banek et al., 2021)