Transcriptomics and Proteomics Methods for *Xenopus* Embryos and Tissues

Michael J. Gilchrist,[1,2,4] Gert Jan C. Veenstra,[1,2,4] and Ken W.Y. Cho[3,4]

[1]*The Francis Crick Institute, London NW1 1AT, United Kingdom;* [2]*Department of Molecular Developmental Biology, Radboud University, 6525 GA Nijmegen, The Netherlands;* [3]*Department of Developmental and Cell Biology, School of Biological Sciences, University of California, Irvine, California 92697*

The general field of quantitative biology has advanced significantly on the back of recent improvements in both sequencing technology and proteomics methods. The development of high-throughput, short-read sequencing has revolutionized RNA-based expression studies, while improvements in proteomics methods have enabled quantitative studies to attain better resolution. Here we introduce methods to undertake global analyses of gene expression through RNA and protein quantification in *Xenopus* embryos and tissues.

OVERVIEW

In recent years, biology has become a more quantitative science, and this trend continues on the back of advances in genomics, proteomics, and analytical methods. *Xenopus*, a well-established model for vertebrate embryology and development, is benefitting from these advances. The availability of more precise data has opened the door to the dissection of intricate biological problems, including protein-interaction networks and gene regulation during cell specification and morphogenesis.

RNA sequencing (RNA-seq), the measurement of gene-specific RNA by counting RNA sequence fragments, is most frequently used to compare gene expression between biological states. Advantages over older methods such as microarrays and quantitative polymerase chain reaction (qPCR) are that it is global, relatively unbiased, and has a very large dynamic range, capable of measuring gene expression with great sensitivity over ∼5 orders of magnitude. RNA-seq requires cellular RNAs to be fragmented into short pieces, which are converted to double-stranded cDNAs. Adaptor sequences are added to the cDNAs and amplified to generate a library, which is then sequenced to generate many millions of reads. These reads are computationally mapped to genes to generate read counts per gene in each sample.

Much of the interest in transcriptomic analysis lies in its ability to compare gene expression, with considerable sensitivity, between two or more states. These states may be developmental stages or tissues; experimental treatments are often compared to controls. To generate statistically meaningful data, experiments need to be carefully designed with sufficient biological replicates. It is also advisable to carefully examine "finished" data sets, looking for consistent expression across groups of similar samples, which may help to identify and eliminate outlying, probably low-quality, samples.

Cite this introduction as *Cold Spring Harb Protoc*; doi:10.1101/pdb.top098350

Alternatively, gene expression can be explored by measuring global protein levels. The drawback is that these experimental methods have lower sensitivity than transcriptomic approaches. Proteomics is, however, a rapidly developing field, and we expect sensitivity to steadily improve.

In the early days of mass spectrometry, the goal was simply to detect proteins, but with increasing emphasis on understanding biological processes, protein quantification in samples is generally required. This can be achieved in tandem mass spectrometry, which fragments sample proteins in two stages for highly refined fragment identification. The first stage generates a precursor ion spectrum of the components obtained after digestion into short peptides; the second stage further fragments these peptides. The combined data from these two stages allow accurate masses to be calculated and hence carefully identify the proteins in the sample. This is, however, not a simple process and unambiguous assignment of all fragments to proteins remains challenging. Nevertheless, quantification is achievable for many proteins, and even posttranslational modifications may be identified where the mass change is sufficient. Quantitative determination of protein levels may be performed in both label-based and label-free methods.

It has recently been shown, but was long suspected, that transcript expression does not correlate well with protein expression in *Xenopus* (Smits et al. 2014; Peshkin et al. 2015). Indeed, this is to be expected in early *Xenopus* development: Many stored maternal transcripts are polyadenylated and translated shortly after fertilization, so the protein levels will lag behind the appearance of the corresponding polyadenylated transcripts. To further complicate matters, there is also a reservoir of maternal proteins, whose half-lives are not known. Combined with the decay rates of the maternal RNAs and their changing polyadenylation status, these factors add layers of complexity to the regulatory events occurring during early embryonic development. Care is clearly needed when extrapolating protein expression levels from mRNA expression levels.

PROTOCOLS

The accompanying protocols provide tools to undertake global analyses of gene expression through RNA and protein quantification in *Xenopus* embryos or tissues. These include generating transcriptomic data via RNA-seq to evaluate the differences in gene expression between biological states, measuring absolute protein levels in biological samples by mass spectrometry, and preparing tissue-specific nuclei (via isolation of nuclei tagged in specific cell types [INTACT]) to analyze the proteomes of selected tissue samples.

High-throughput sequencing can be carried out on RNA samples from embryos or adult tissues to generate counts per gene per sample. These data can then be used to investigate differences in gene expression levels between states, as described in Protocol: **An RNA-Seq Protocol for Differential Expression Analysis** (Owens et al. 2019). Similar approaches have been widely used in the *Xenopus* community: Examples include analyzing the role of Sox7 in *Xenopus* germ cell development by studying differential expression between primordial germ cells and endoderm cells (Butler et al. 2018); examining the role of maternal transcription factors such as Foxh1, VegT, and Otx1 in zygotic genome activation by injecting translation-blocking antisense morpholino oligonucleotides (Charney et al. 2017; Paraiso et al. 2019); and elucidating different signatures for the development of dorsal and ventral lips of the early *Xenopus* gastrula embryo using gene set expression analysis (GSEA) (Ding et al. 2017).

To circumvent the poor correlation between mRNA and protein levels in early stage *Xenopus* embryos, gene expression dynamics can be studied though proteome profiling, as described in Protocol: **Mass Spectrometry-Based Absolute Quantification of Single *Xenopus* Embryo Proteomes** (Lindeboom et al. 2019). Although care must be taken to work around the yolk proteins in *Xenopus*, the large eggs and cells from the early embryonic divisions contain sufficient protein (unlike those in early mammalian development) to enable quantification of thousands of proteins using mass spectrometry-based proteomics.

There have been steady developments in quantitative proteomics in *Xenopus*, especially from the Kirschner lab, which described protein expression dynamics of whole embryos during embryogenesis

(Wühr et al. 2014; Presler et al. 2017), and from the Nemes and Moody labs, which studied the proteomics of single blastomeres (Onjiko et al. 2015; Lombard-Banek et al. 2016). Together, these highlight the important contribution of the *Xenopus* system in probing early development from a protein perspective.

Pure cell populations may be isolated from embryonic tissues by tagging the nuclei of specific cell types, which are then used for the proteomic analysis of nuclear protein complexes, as described in Protocol: **INTACT Proteomics in *Xenopus*** (Wasson et al. 2019). This biotin-based assay is used to separate the nuclei of interest from those of other cell types without requiring specific antibodies, although it does require transgenic animals. Coupled with liquid chromatography tandem mass spectrometry, this approach allowed isolation of nuclei from developing *Xenopus* heart cells and has identified *Xenopus* orthologs of human genes implicated in congenital heart diseases (Amin et al. 2014).

FUTURE CONSIDERATIONS

The specific properties of the *Xenopus* system have encouraged time-resolved experiments in early development, enabling studies of the dynamics (Collart et al. 2014) and absolute quantitation (Owens et al. 2016) of gene expression. Steadily improving low-RNA capabilities of sequencing protocols have enabled the spatial resolution of gene expression between axially mapped blastomeres of the eight-cell embryo (De Domenico et al. 2015) and different regions of gastrula stage embryos (Blitz et al. 2017).

Single-cell RNA-seq developed using droplet microfluidics (Klein et al. 2015) has enabled the temporal and spatial analysis of gene expression in *Xenopus* embryos, producing a catalog of developmental cell states and a map of cell lineage differentiation spanning the blastula stage to early organogenesis (Briggs et al. 2018). We also anticipate that new technologies, such as RNA seqFISH+ (Eng et al. 2019), may be used to study the spatial distribution of transcripts in the embryo, which is not possible in current single-cell RNA-seq approaches.

The other significant development in sequencing technology is that of long-read, single-molecule sequencing from the likes of PacBio (McCarthy 2010) and Oxford Nanopore (Laver et al. 2015). The short fragments used in Illumina sequencing do not allow exon connectivity to be determined over whole transcripts, and these new long reads provide the perfect tool to properly explore transcript isoform diversity. This is very relevant to developing our understanding of tissue-specific alternative splicing. The relatively small numbers of reads generated in long-read sequencing largely preclude useful quantitation, although these technologies are steadily improving. Short-read sequencing is currently the preferred choice for quantitative biology, but long-read sequencing has the potential to revolutionize our understanding of haplotypes at the genome level and exon usage at the transcript level.

It is inevitable that these trends will come together in single-cell proteomics, and that the resolving power of such techniques will steadily improve. Already there is a study that examines the coexpression of lineage-specific transcription factors in individual cells during human erythropoiesis (Palii et al. 2019). There is no doubt that *Xenopus* will continue to make a powerful contribution to biomedicine, as the lens of biological science, while retaining the global viewpoint, focuses on ever smaller elements of the system.

REFERENCES

Amin NM, Greco TM, Kuchenbrod LM, Rigney MM, Chung MI, Wallingford JB, Cristea IM, Conlon FL. 2014. Proteomic profiling of cardiac tissue by isolation of nuclei tagged in specific cell types (INTACT). *Development* **141:** 962–973. doi:10.1242/dev.098327

Blitz IL, Paraiso KD, Patrushev I, Chiu WTY, Cho KWY, Gilchrist MJ. 2017. A catalog of *Xenopus tropicalis* transcription factors and their regional expression in the early gastrula stage embryo. *Dev Biol* **426:** 409–417. doi:10.1016/j.ydbio.2016.07.002

Briggs JA, Weinreb C, Wagner DE, Megason S, Peshkin L, Kirschner MW, Klein AM. 2018. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360:** eaar5780. doi:10.1126/science.aar5780

M. J. Gilchrist et al.

Butler AM, Owens DA, Wang L, King ML. 2018. A novel role for *sox7* in *Xenopus* early primordial germ cell development: mining the PGC transcriptome. *Development* **145**: dev155978. doi:10.1242/dev.155978

Charney RM, Forouzmand E, Cho JS, Cheung J, Paraiso KD, Yasuoka Y, Takahashi S, Taira M, Blitz IL, Xie X, et al. 2017. Foxh1 occupies cis-regulatory modules prior to dynamic transcription factor interactions controlling the mesendoderm gene program. *Dev Cell* **40**: 595–607 e594. doi:10.1016/j.devcel.2017.02.017

Collart C, Owens ND, Bhaw-Rosun L, Cooper B, De Domenico E, Patrushev I, Sesay AK, Smith JN, Smith JC, Gilchrist MJ. 2014. High-resolution analysis of gene activity during the *Xenopus* mid-blastula transition. *Development* **141**: 1927–1939. doi:10.1242/dev.102012

De Domenico E, Owens ND, Grant IM, Gomes-Faria R, Gilchrist MJ. 2015. Molecular asymmetry in the 8-cell stage *Xenopus tropicalis* embryo described by single blastomere transcript sequencing. *Dev Biol* **408**: 252–268. doi:10.1016/j.ydbio.2015.06.010

Ding Y, Colozza G, Zhang K, Moriyama Y, Ploper D, Sosa EA, Benitez MDJ, De Robertis EM. 2017. Genome-wide analysis of dorsal and ventral transcriptomes of the *Xenopus laevis* gastrula. *Dev Biol* **426**: 176–187. doi:10.1016/j.ydbio.2016.02.032

Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, Yun J, Cronin C, Karp C, Yuan GC, et al. 2019. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**: 235–239. doi:10.1038/s41586-019-1049-y

Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW. 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**: 1187–1201. doi:10.1016/j.cell.2015.04.044

Laver T, Harrison J, O'Neill PA, Moore K, Farbos A, Paszkiewicz K, Studholme DJ. 2015. Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomol Detect Quantif* **3**: 1–8. doi:10.1016/j.bdq.2015.02.001

Lindeboom RGH, Smits AH, Perino M, Veenstra GJC, Vermeulen M. 2019. Mass spectrometry-based absolute quantification of single *Xenopus* embryo proteomes. *Cold Spring Harb Protoc* **2019**: pdb.prot098376. doi:10.1101/pdb.prot098376

Lombard-Banek C, Reddy S, Moody SA, Nemes P. 2016. 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. doi:10.1074/mcp.M115.057760

McCarthy A. 2010. Third generation DNA sequencing: Pacific Biosciences' single molecule real time technology. *Chem Biol* **17**: 675–676. doi:10.1016/j.chembiol.2010.07.004

Onjiko RM, Moody SA, Nemes P. 2015. Single-cell mass spectrometry reveals small molecules that affect cell fates in the 16-cell embryo. *Proc Natl Acad Sci* **112**: 6545–6550. doi:10.1073/pnas.1423682112

Owens NDL, Blitz IL, Lane MA, Patrushev I, Overton JD, Gilchrist MJ, Cho KWY, Khokha MK. 2016. Measuring absolute RNA copy numbers at high temporal resolution reveals transcriptome kinetics in development. *Cell Rep* **14**: 632–647. doi:10.1016/j.celrep.2015.12.050

Owens NDL, De Domenico E, Gilchrist MJ. 2019. An RNA-seq protocol for differential expression analysis. *Cold Spring Harb Protoc* **2019**: pdb.prot098368. doi:10.1101/pdb.prot098368

Palii CG, Cheng Q, Gillespie MA, Shannon P, Mazurczyk M, Napolitani G, Price ND, Ranish JA, Morrissey E, Higgs DR, et al. 2019. Single-cell proteomics reveal that quantitative changes in co-expressed lineage-specific transcription factors determine cell fate. *Cell Stem Cell* **24**: 812–820 e815. doi:10.1016/j.stem.2019.02.006

Paraiso KD, Blitz IL, Coley M, Cheung J, Sudou N, Taira M, Cho KWY. 2019. Endodermal maternal transcription factors establish super-enhancers during zygotic genome activation. *Cell Rep* **27**: 2962–2977 e2965. doi:10.1016/j.celrep.2019.05.013

Peshkin L, Wühr M, Pearl E, Haas W, Freeman RM Jr, Gerhart JC, Klein AM, Horb M, Gygi SP, Kirschner MW. 2015. On the relationship of protein and mRNA dynamics in vertebrate embryonic development. *Dev Cell* **35**: 383–394. doi:10.1016/j.devcel.2015.10.010

Presler M, Van Itallie E, Klein AM, Kunz R, Coughlin ML, Peshkin L, Gygi SP, Wühr M, Kirschner MW. 2017. Proteomics of phosphorylation and protein dynamics during fertilization and meiotic exit in the *Xenopus* egg. *Proc Natl Acad Sci* **114**: E10838–E10847. doi:10.1073/pnas.1709207114

Smits AH, Lindeboom RG, Perino M, van Heeringen SJ, Veenstra GJ, Vermeulen M. 2014. Global absolute quantification reveals tight regulation of protein expression in single *Xenopus* eggs. *Nucleic Acids Res* **42**: 9880–9891. doi:10.1093/nar/gku661

Wasson L, Amin NM, Conlon FL. 2019. INTACT proteomics in *Xenopus*. *Cold Spring Harb Protoc* **2019**: pdb.prot098384. doi:10.1101/pdb.prot098384

Wühr M, Freeman RM Jr, Presler M, Horb ME, Peshkin L, Gygi S, Kirschner MW. 2014. Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. *Curr Biol* **24**: 1467–1475. doi:10.1016/j.cub.2014.05.044

Cite this introduction as *Cold Spring Harb Protoc;* doi:10.1101/pdb.top098350

Cold Spring Harbor Protocols

Transcriptomics and Proteomics Methods for *Xenopus* Embryos and Tissues

Michael J. Gilchrist, Gert Jan C. Veenstra and Ken W.Y. Cho

Email Alerting Service	Receive free email alerts when new articles cite this article - **click here.**
Subject Categories	Browse articles on similar topics from *Cold Spring Harbor Protocols.* Analysis of Gene Expression (187 articles) Analysis of Gene Expression, general (128 articles) High-Throughput Analysis, general (150 articles) Mass Spectrometry (78 articles) Proteins and Proteomics, general (557 articles) RNA (297 articles) RNA, general (251 articles) Xenopus (158 articles)

To subscribe to *Cold Spring Harbor Protocols* go to:
http://cshprotocols.cshlp.org/subscriptions