



Published in final edited form as:

*Cold Spring Harb Protoc.* ; 2020(5): pdb.top097915. doi:10.1101/pdb.top097915.

## Genomics Methods for *Xenopus* Embryos and Tissues

Michael J. Gilchrist<sup>1,4</sup>, Ken W.Y. Cho<sup>2,4</sup>, Gert Jan C. Veenstra<sup>3,4</sup>

<sup>1</sup>The Francis Crick Institute, London NW1 1AT, United Kingdom; <sup>2</sup>Department of Developmental and Cell Biology, University of California, Irvine, California 92697 <sup>3</sup>Radboud University, Department of Molecular Developmental Biology, 6525GA Nijmegen, The Netherlands

### Abstract

High-throughput sequencing methods have created exciting opportunities to explore the regulatory landscape of the entire genome. Here we introduce methods to characterize the genomic locations of bound proteins, open chromatin, and sites of DNA–DNA contact in *Xenopus* embryos. These methods include chromatin immunoprecipitation followed by sequencing (ChIP-seq), a combination of DNase I digestion and sequencing (DNase-seq), the assay for transposase-accessible chromatin and sequencing (ATAC-seq), and the use of proximity-based DNA ligation followed by sequencing (Hi-C).

### OVERVIEW

The epigenetic state of chromatin regulates gene expression—and hence cellular differentiation—by controlling the access of transcription factors to DNA. Here we introduce methods to explore the locations of bound proteins, open chromatin, and sites of DNA–DNA contact at the whole-genome level in *Xenopus* embryos and tissues. These methods of data generation are often used in conjunction with gene expression analyses, which we discuss in Introduction: Transcriptomics and Proteomics Methods for *Xenopus* Embryos and Tissues (Gilchrist et al. 2019).

Chromatin immunoprecipitation (ChIP) is one of the most direct ways to identify the sites of interaction between the genome and DNA-binding proteins (typically transcription factors and histones). In its most straightforward application, ChIP relies on an initial cross-linking stage in the intact animals or dissected tissues, followed by DNA fragmentation by sonication. Complexes containing DNA fragments bound to the protein of interest are immunoprecipitated with an antibody that recognizes the protein. After dissolution of the cross-linking, the DNA fragments are recovered and subjected to direct high-throughput sequencing (hence the abbreviation ChIP-seq).

Mapping regions of relatively open chromatin is valuable for identifying regulatory elements, including promoters and enhancers. Even when bound by proteins such as transcription factors, the DNA of the regulatory element remains accessible, rendering these

<sup>4</sup>Correspondence: drmikeygilchrist@gmail.com; kwcho@uci.edu; g.veenstra@science.ru.nl.

regions relatively sensitive to cleavage by DNaseI or Tn5 transposase. Two methods, DNase-seq (see Protocol: DNase-seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells [Song and Crawford 2010]) and the assay for transposase-accessible chromatin (ATAC) and sequencing (ATAC-seq) (Buenrostro et al. 2013), exploit this feature to map regions of accessible DNA. In these methods, digested double-cut fragments of DNA are either directly amplified (as in ATAC-seq) or amplified after purification (as in DNase-seq) and then sequenced. These approaches provide the ability to discover important new enhancers without having a priori knowledge of the location and identity of bound transcription factors.

In a rather different approach to genomic data generation, we can determine the distribution of DNA–DNA contacts within and between chromosomes to understand how chromosomal DNA is folded in the nucleus. Cross-linking is used to form bridges at the contact points, and after DNA fragmentation, the biotinylated loose ends of DNA fragments in the same complex are ligated. After removal of the cross-links and further digestion, the joined fragments are isolated and sequenced from both ends to identify the different genomic regions that were in contact. This global and relatively unbiased variant of the chromosome conformation capture methods is referred to as Hi-C. The sequence data contain information on the DNA looping structures understood to regulate transcription, and are thus rather different from, but highly complementary to, the data generated by ChIP-seq, DNase-seq, and ATAC-seq experiments.

In general, these experimental methods are not very efficient, requiring relatively large numbers of cells to provide a robust signal. The *Xenopus* system is therefore ideal for these types of applications because of the ability to collect large numbers of synchronously developing embryos after in vitro fertilization. Below we introduce ChIP-seq, DNase-seq, ATAC-seq, and Hi-C protocols that have been developed specifically for use in *Xenopus* embryos and tissues and provide examples of their applications.

The *Xenopus* community is fortunate in having two well-assembled genomes, that of the diploid *X. tropicalis* (Hellsten et al. 2010) and that of the allo-tetraploid *X. laevis* (Session et al. 2016). This is important, as the sequence fragments generated by these methods generally map to the intergenic and intragenic noncoding (introns) regions, and incomplete assemblies will cause loss of potentially valuable data. The diploid genome of *X. tropicalis* makes the high-throughput genomic data simpler to interpret when compared to data from the larger, partly duplicated genome of *X. laevis*. However, the larger *X. laevis* embryos may be preferred for ease of experimental manipulation.

## PROTOCOLS

Two ChIP-seq protocols, Protocol: Mapping Chromatin Features of *Xenopus* Embryos (Gentsch and Smith 2019) and Protocol: ChIP-Sequencing in *Xenopus* Embryos (Hontelez et al. 2019), offer slightly different approaches to prepare enriched nuclei and yolk-depleted embryo lysate. In particular, ChIP with cleared lysates—as described in the latter protocol—requires less starting material. ChIP-seq has been widely used in the *Xenopus* community to understand the targets of developmentally important transcription factors—for example,

Smad2/3 (Yoon et al. 2011); Foxh1 (Chiu et al. 2014); T-Box family proteins (Gentsch et al. 2013);  $\beta$ -catenin (Nakamura et al. 2016); Otx2, Lim1/Lhx1, and Gsc (Yasuoka et al. 2014); Vegt and Otx1 in early embryos (Paraiso et al. 2019); and Prdm12 in developing inhibitory neurons important for vertebrate locomotion (Thélie et al. 2015). ChIP-seq has also been used to identify promoters and compare them across species (van Heeringen et al. 2011) or to look at the dynamic changes in genome-wide distribution of histone modifications and RNA polymerase II occupancy (Akkers et al. 2009; Hontelez et al. 2015).

Open chromatin is generally considered a prerequisite for binding of transcription factors, and two protocols can be used to study the distribution of open chromatin regions over the *Xenopus* genome: One uses DNase-seq (Protocol: DNase-seq to Study Chromatin Accessibility in Early *Xenopus tropicalis* Embryos [Cho et al. 2019]), and the other uses ATAC-seq (Protocol: Assay for Transposase-Accessible Chromatin-Sequencing Using *Xenopus* Embryos [Bright and Veenstra 2019]). Reads from these protocols are mapped to the genome and generally produce peaks or otherwise-delineated small regions of DNA. These regions can then be analyzed for enrichment in DNA-binding motifs to predict which transcription factors may be active in particular developmental stages or tissues. If enough sequence reads are mapped, specific transcription factor footprints can be resolved at single-base resolution to determine how those factors contact the DNA (Boyle et al. 2011; Neph et al. 2012; Buenrostro et al. 2013).

Most chromosome conformation capture methods use defined viewpoints (genomic regions of interest for which interactions are determined) to build dense maps of DNA–DNA contact information (for review, see Nicoletti et al. 2018). However, it is also possible to use this approach to perform an unbiased analysis in an organism's cells to study the three-dimensional organization of chromosomes under physiological conditions. This approach is described for *Xenopus* in Protocol: Generating a Three-Dimensional Genome from *Xenopus* with Hi-C (Quigley and Heinz 2019). This approach has previously been used in *Xenopus* to investigate the relationship between Foxj1 binding and chromatin loops in multiciliated cells (Quigley and Kintner 2017) using tethered conformation capture (Kalhor et al. 2011).

## FUTURE CONSIDERATIONS

The major change sweeping through the world of biological data generation is the switch in emphasis from the bulk analysis of whole organisms or tissues to the analysis of hundreds to many thousands of individual cells from an embryo or tissue sample. Advances have been most rapid in the field of single cell transcriptomics - with, for example, the recent launch of the Human Cell Atlas (Rozenblatt-Rosen et al. 2017), which aims to create comprehensive reference maps of cell types in all major tissues. Recently, a major single-cell transcriptomics survey was conducted in *Xenopus* specimens spanning the blastula stage to the tailbud stage; this enabled the characterization of cell types from the earliest pluripotent cells to the well-differentiated cells of early organogenesis (Briggs et al. 2018). Although not all mRNAs in each cell are currently detected, single-cell RNA sequencing (scRNA-seq) can help us define the set of cell types in a population of cells, determine the evolution of cell lineages during development, and better understand the regulatory relationships between genes. Inevitably, the very large data sets that are generated do introduce some challenges in

the development and application of computational methods, but modern approaches such as deep learning may be ideal for this.

Genomics techniques will not be far behind. Already, a number of approaches to single-cell genomics analysis are being developed along the lines of the multicellular methods, although the relatively small amounts of DNA in each sample will require highly efficient experimental methods and some rethinking of the downstream analysis. New approaches include the use of the Tn5 transposase for efficient library preparation—for example, in the CUT&Tag method (Kaya-Okur et al. 2019) for efficient epigenomic profiling on low cell numbers and single cells. An ATAC-seq method has been developed for single-cell work (Chen et al. 2018) and used to resolve dynamic changes in the chromatin landscape and to uncover the *cis*-regulatory programs of *Drosophila* germ layer formation (Cusanovich et al. 2018). Last, in addition to Hi-C, Capture Hi-C (Jäger et al. 2015) has been developed to specifically enrich (for example) for promoter-containing fragments from Hi-C libraries, and it will be able to generate evidence for dynamic interactions between promoters and enhancers.

The availability of genome assemblies for two closely related *Xenopus* species (see above) provides opportunities for new insights in comparative biology. In addition, there are many potential sources of genetic variation data captured in expressed sequences from both species and from different strains of these. Analysis of such data is useful for many different purposes, such as the design of morpholinos and CRISPR guide RNAs when using outbred populations. Genetic variation is also key to analyses of quantitative trait loci (QTL) and genome evolution. Although analysis of variation in whole-genome sequencing data is outside the scope of this introduction, genomic variation in *X. laevis* and *X. tropicalis* has already been analyzed in several reports (Elurbe et al. 2017; Savova et al. 2017; Mitros et al. 2019). Further studies will likely address genetic variation and their associated developmental, gene-regulatory, and phenotypic traits, both within populations and between closely related species.

Future genome-wide studies in *Xenopus* using the techniques described above will lead to the discovery of new mechanisms and to a better understanding of the processes controlling transcription and gene activation. We look forward to seeing the *Xenopus* community employing this next generation of methods in this most tractable of model systems.

## REFERENCES

1. Akkers RC, van Heeringen SJ, Jacobi UG, Janssen Megens EM, François KJ, Stunnenberg HG, Veenstra GJ. 2009 A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell* 17: 425–434. doi:10.1016/j.devcel.2009.08.005 [PubMed: 19758566]
2. Boyle AP, Song L, Lee BK, London D, Keefe D, Birney E, Iyer VR, Crawford GE, Furey TS. 2011 High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Res* 21: 456–464. doi:10.1101/gr.112656.110 [PubMed: 21106903]
3. Briggs JA, Weinre 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 [PubMed: 29700227]
4. Bright AR, Veenstra GJC. 2019 Assay for transposase-accessible chromatin-sequencing using *Xenopus* embryos. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot098327.

5. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013 Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10: 1213–1218. doi:10.1038/nmeth.2688 [PubMed: 24097267]
6. Chen X, Miragaia RJ, Natarajan KN, Teichmann SA. 2018 A rapid and robust method for single cell chromatin accessibility profiling. *Nat Commun* 9: 5345. doi:10.1038/s41467-018-07771-0 [PubMed: 30559361]
7. Chiu WT, CharneyLe R, Blitz IL, Fish MB, Li Y, Biesinger J, Xie X, Cho KW. 2014 Genome-wide view of TGF $\beta$ /Foxh1 regulation of the early mesendoderm program. *Development* 141: 4537–4547. doi:10.1242/dev.107227 [PubMed: 25359723]
8. Cho JS, Blitz IL, Cho KWY. 2019 DNase-seq to study chromatin accessibility in early *Xenopus tropicalis* embryos. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot098335
9. Cusanovich DA, Reddington JP, Garfield DA, Daza RM, Aghamirzaie D, Marco-Ferreres R, Pliner HA, Christiansen L, Qiu X, Steemers FJ, et al. 2018 The *cis*-regulatory dynamics of embryonic development at single-cell resolution. *Nature* 555: 538–542. doi:10.1038/nature25981 [PubMed: 29539636]
10. Elurbe DM, Paranjpe SS, Georgiou G, van Kruijsbergen I, Bogdanovic O, Gibeaux R, Heald R, Lister R, Huynen MA, van Heeringen SJ, et al. 2017 Regulatory remodeling in the allo-tetraploid frog *Xenopus laevis*. *Genome Biol* 18: 198. doi:10.1186/s13059-017-1335-7 [PubMed: 29065907]
11. Gentsch GE, Smith JC. 2019 Mapping chromatin features of *Xenopus* embryos. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot100263
12. Gentsch GE, Owens ND, Martin SR, Piccinelli P, Faial T, Trotter MW, Gilchrist MJ, Smith JC. 2013 In vivo T-box transcription factor profiling reveals joint regulation of embryonic neuromesodermal bipotency. *Cell Rep* 4: 1185–1196. doi:10.1016/j.celrep.2013.08.012 [PubMed: 24055059]
13. Gilchrist MJ, Veenstra GJC, Cho KWY. 2019 Transcriptomics and proteomics methods for *Xenopus* embryos and tissues. *Cold Spring Harb Protoc* doi:10.1101/pdb.top098350
14. Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I, Putnam NH, Shu S, Taher L, et al. 2010 The genome of the western clawed frog *Xenopus tropicalis*. *Science* 328: 633–636. doi:10.1126/science.1183670 [PubMed: 20431018]
15. Hontelez S, van Kruijsbergen I, Georgiou G, van Heeringen SJ, Bogdanovic O, Lister R, Veenstra GJC. 2015 Embryonic transcription is controlled by maternally defined chromatin state. *Nat Commun* 6: 10148. doi:10.1038/ncomms10148 [PubMed: 26679111]
16. Hontelez S, van Kruijsbergen I, Veenstra GJC. 2019 ChIP-sequencing in *Xenopus* embryos. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot097907
17. Jäge R, Migliorini G, Henrion M, Kandaswamy R, Speedy HE, Heindl A, Whiffin N, Carnicer MJ, Broome L, Dryden N, et al. 2015 Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat Commun* 6: 6178. doi:10.1038/ncomms7178 [PubMed: 25695508]
18. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L. 2011 Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat Biotechnol* 30: 90–98. doi:10.1038/nbt.2057 [PubMed: 22198700]
19. Kaya Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. 2019 CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* 10: 1930. doi:10.1038/s41467-019-09982 [PubMed: 31036827]
20. Mitros T, Lyons JB, Sessio AM, Jenkins J, Shu S, Kwon T, Lane M, Ng C, Grammer TC, Khokha MK, et al. 2019 A chromosome-scale genome assembly and dense genetic map for *Xenopus tropicalis*. *Dev Biol* 452: 8–20. doi:10.1016/j.ydbio.2019.03.015 [PubMed: 30980799]
21. Nakamura Y, de Paiva Alves E, Veenstra GJ, Hoppler S. 2016 Tissue- and stage-specific Wnt target gene expression is controlled subsequent to beta-catenin recruitment to *cis*-regulatory modules. *Development* 143: 1914–1925. doi:10.1242/dev.131664 [PubMed: 27068107]
22. Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E, Vernot B, Thurman RE, John S, Sandstrom R, Johnson AK, et al. 2012 An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489: 83–90. doi:10.1038/nature11212 [PubMed: 22955618]

23. Nicoletti C, Forcato M, Bicciato S. 2018 Computational methods for analyzing genome-wide chromosome conformation capture data. *Curr Opin Biotechnol* 54: 9–105. doi:10.1016/j.copbio.2018.01.023
24. 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 [PubMed: 31167141]
25. Quigley IK, Heinz S. 2019 Generating a three-dimensional genome from *Xenopus* with Hi-C. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot098343
26. Quigley IK, Kintner C. 2017 Rfx2 stabilizes Foxj1 binding at chromatin loops to enable multiciliated cell gene expression. *PLoS Genet* 13: e1006538. doi:10.1371/journal.pgen.1006538 [PubMed: 28103240]
27. Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA. 2017 The human cell atlas: from vision to reality. *Nature* 550: 451–453. doi:10.1038/550451a [PubMed: 29072289]
28. Savova V, Pearl EJ, Boke E, Nag A, Adzhubei I, Horb ME, Peshkin L. 2017 Transcriptomic insights into genetic diversity of protein-coding genes in *X. laevis*. *Dev Biol* 424: 181–188. doi:10.1016/j.ydbio.2017.02.019 [PubMed: 28283406]
29. Session AM, Uno Y, Kwon T, Chapman JA, Toyoda A, Takahashi S, Fukui A, Hikosaka A, Suzuki A, Kondo M, et al. 2016 Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* 538: 336–343. doi:10.1038/nature19840 [PubMed: 27762356]
30. Song L, Crawford GE. 2010 DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot5384
31. Th  lie A, Desiderio S, Hanotel J, Quigley I, Van Driessche B, Rodari A, Borromeo MD, Kricha S, Lahaye F, Croce J, et al. 2015 Prdm12 specifies V1 interneurons through cross-repressive interactions with Dbx1 and Nkx6 genes in *Xenopus*. *Development* 142: 3416–3428. doi:10.1242/dev.121871 [PubMed: 26443638]
32. Van Heeringen SJ, Akhtar W, Jacobi UG, Akkers RC, Suzuki Y, Veenstra GJ. 2011 Nucleotide composition-linked divergence of vertebrate core promoter architecture. *Genome Res* 21: 410–421. doi:10.1101/gr.111724.110 [PubMed: 21284373]
33. Yasuoka Y, Suzuki Y, Takahashi S, Someya H, Sudou N, Haramoto Y, Cho KWY, Asashima M, Sugano S, Taira M. 2014 Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification. *Nat Commun* 5: 4322. doi:10.1038/ncomms5322
34. Yoon SJ, Wills AE, Chuong E, Gupta R, Baker JC. 2011 HEB and E2A function as SMAD/FOXH1 cofactors. *Genes Dev* 25: 1654–1661. doi:10.1101/gad.16800511 [PubMed: 21828274]