



MILESTONE 4

ChIP-seq captures the chromatin landscape

The fundamental instructions for cellular function are encoded in DNA, which tightly associates with histones to form a complex structure called chromatin. This ensures the stability, replication and proper interpretation of the genetic code.

Chromatin is physically linked to and regulated by a multitude of proteins (for example, transcription factors) and RNAs, ensuring that genes are correctly expressed or silenced in the appropriate cellular context. To understand how chromatin-bound proteins affect gene expression and, ultimately, cell behaviour, it is essential to characterize how, where and when these regulatory proteins bind to chromatin.

Chromatin immunoprecipitation, a technique that goes back to the 1980s, allows the identification of DNA regions that are bound by a protein of interest. First, cells are treated with one or more crosslinking agents (usually formaldehyde) so that covalent bonds are formed between DNA and associated proteins, thus preserving structural and regulatory interactions. Then, the chromatin is fragmented, and an antibody that recognizes a specific protein is used to capture its DNA-bound fragments. Finally, the induced covalent bonds are broken, and the DNA is purified for further analysis.

Initially, these immunoprecipitated DNA fragments were

hybridized to microarray platforms (as in the 'ChIP-chip' method) to gain a genome-scale perspective. However, this assay has substantial limitations, including restricted resolution and genome coverage and noisy signals.

In 2007, capitalizing on the rise of next-generation sequencing technologies, ChIP-seq was born, and a series of papers drafted the first genome-wide landscapes of protein-DNA interactions at high resolution. Early studies focused on histone post-translational modifications (Barski et al. and Mikkelsen et al.) and transcription factors (Johnson et al. and Robertson et al.).

Large initiatives such as the ENCODE (MILESTONE 14) or Roadmap Epigenomics Mapping consortia were among the pioneers that leveraged ChIP-seq to characterize the epigenomic profiles of a variety of broadly used cell lines and primary cell types and tissues. To this day, these maps constitute reference datasets for the research community.

However, ChIP-seq conducted on heterogeneous cell populations can mask phenomena unique to or more prevalent in certain subpopulations. Technological advances have enabled the development of single-cell ChIP-seq (Rotem et al. and Grosselin et al.). Although these techniques represent exciting steps forwards, they remain technically challenging. Recent and promising alternatives to ChIP-seq include CUT&RUN (Skene and Henikoff) and CUT&Tag sequencing (Kaya-Okur et al.). The advantages over ChIP-seq include not requiring crosslinking and providing a high signal-to-noise ratio at lower sequencing depth. Nevertheless, ChIP-seq continues to be a standard method widely used in transcriptional and epigenetic studies.

The rapid and widespread adoption of ChIP-seq by researchers across fields, such as development, evolution and cancer, provided the basis for a notable leap in our understanding of chromatin biology. By carefully studying patterns of histone modifications and transcription factor binding with improved resolution and with the aid of sophisticated computational tools, scientists deepened their knowledge about basic gene regulatory mechanisms and the role of non-coding genetic variants in disease.

Despite the current popularity of ChIP-seq, available transcription factor and histone mark ChIP-seq data in different cell contexts remain sparse. Data integration and imputation methods using published ChIP-seq are likely to contribute significantly to the ongoing quest to decipher gene regulatory mechanisms in physiological processes and diseases.

The biological insights obtained in the last decade thanks to the use of ChIP-seq data have been and continue to be transformational.

Tiago Faial, *Nature Genetics*

“ ChIP-seq ... provided the basis for a notable leap in our understanding of chromatin biology ”

ORIGINAL ARTICLES Barski, A. et al. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007) | Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007) | Johnson, D. S. et al. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497–1502 (2007) | Robertson, G. et al. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657 (2007)
FURTHER READING Rotem, A. et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat. Biotechnol.* **33**, 1165–1172 (2015) | Grosselin, K. et al. High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat. Genet.* **51**, 1060–1066 (2019) | Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856 (2017) | Kaya-Okur, H. S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019)