MILESTONE 10

Probing nuclear architecture with Hi-C

Hi-C revolutionized our view of the arrangement of DNA in the nucleus. Before, most insights into nuclear packing came from light microscopy, for which specific loci were tagged with fluorescent proteins to measure their spatial relationship. Here, the major limitation is that the resolution is too low to probe interactions at fine-grained base pair level. Analysis is limited to only a few loci at a time, and cell line construction can be laborious, so the process is low-throughput.

In 2002, Dekker et al. published the chromatin conformation capture (3C) technique, where DNA is crosslinked to its supporting proteins to fix its conformation in space, then digested with a restriction enzyme, leaving sticky ends that can be re-ligated. Any two sequences close in 3D space, but not necessarily close on the linear DNA molecule, can be joined. After crosslink reversal and DNA fragmenting, the hybrid DNA molecules created by ligation of distal sequences can be identified by PCR using primers that bind to each fragment. However, this low-throughput, targeted approach does not enable the discovery of novel interactions.

> “a real whole-genome view of higher-order chromatin architecture”

Credit: Marcio Silva / Alamy Stock Photo

Modifications to this method were made to improve the number of interactions that could be measured, leading to chromatin conformation capture on a chip (4C) and chromatin conformation capture carbon copied (5C). In 2009, Lieberman-Aiden et al. published the Hi-C approach, which used high-throughput short-read sequencing to sequence all hybrid DNA fragments after the ligation step. It was thus possible to identify nearly all fragments of the genome that are physically adjacent to all others, giving a real whole-genome view of higher-order chromatin architecture. For any given nucleus in the sample, the crosslinking provides a snapshot of the interactions at that time. In a bulk sample, the number of reads for a given ligation product increases the more often those two fragments are physically (or spatially) proximal to each other.

The first Hi-C study identified the existence of two nuclear compartments. Sequences in one compartment are more likely to interact with each other, and interactions between the compartments are rare. The compartments were arbitrarily labelled A and B, with the A compartment containing mostly open chromatin with expressed genes and the B compartment containing more densely packed heterochromatin.

Subsequent higher-resolution studies revealed the existence of topologically associating domains (TADs) in mammalian cells and *Drosophila*. TADs are observed at a finer scale than A and B compartments, and intra-TAD sequences interact more frequently than inter-TAD sequences. Domain boundaries are marked by various features including binding sites for CTCF and cohesin proteins.

Technical advances revealed finer-scale structure; sub-TADs and loops are nested within TADs and differ between differentiated and undifferentiated cells, with some changes observed to be correlated with gene expression changes.

Hi-C represents an averaged view of interaction frequency across a population of crosslinked cells; the extent to which a given interaction occurred in individual cells was not clear. The first single-cell Hi-C study found that single cells have similar domain structure to that seen in bulk Hi-C but that there is stochastic variation in interdomain contacts between cells. Further single-cell studies gave a more dynamic view of chromosome movements and showed that cells could be grouped according to cell cycle stage.

Hi-C has also been used in chromosome scaffolding pipelines for genome assembly to order sequencing reads (MILESTONE 16).

Since the first Hi-C study, our knowledge of how chromosomes are arranged in the nucleus has improved immeasurably, but there is still much to learn. In particular, the functional consequences of the various levels of organization, and how they relate to transcription, are still to be determined. Further refinements and improvements to the Hi-C method will undoubtedly contribute to this understanding in the coming years.

Andrew Cosgrove, *Genome Biology*

ORIGINAL ARTICLE Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009)
FURTHER READING Dekker, J. et al. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002) | Zhao, Z. et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* **38**, 1341–1347 (2006) | Simonis, M. et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nat. Genet.* **38**, 1348–1354 (2006) | Dostie, J. et al. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006) | Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012) | Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012) | Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012) | Phillips-Cremins, J. E. et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281–1295 (2013) | Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014) | Nagano, T. et al. Single cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013) | Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017)