GENOMICS

Toward a 3D genome in high resolution

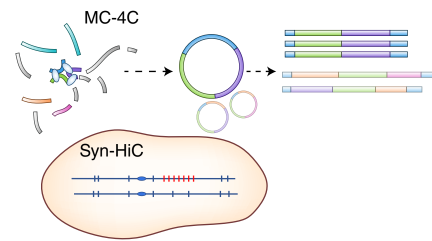
Independent efforts shine light on the 3D genome structure by looking at multiple contacts along an allele or equalizing the distance between restriction sites for higher-resolution Hi-C maps.

Hidden in the 3D structure of the genome are answers to many questions about processes from transcriptional regulation to genome stability. Striving to understand how exactly the long DNA polymer is folded into the small nucleus, many researchers are developing ever more intricate methods, but several formidable challenges remain. Two of these were recently tackled from two very different angles, from groups in the Netherlands and France, with the similar goal of improving the resolution of genomic interaction maps.

Current molecular methods such as the suite of chromosome conformation capture (3C) tools mainly look at pairwise interactions of DNA, but there are probably many more points of interaction along a DNA molecule that define the small- and large-scale architecture of the genome. Wouter de Laat from the Hubrecht Institute and Jeroen de Ridder from the University Medical Center in Utrecht, along with their respective PhD students Carlo Vermeulen and Amin Allahyar, wanted to explore higher-order interactions and their biological implications.

In what de Ridder calls a "beautiful collaboration between computational and genomic teams," they modified 4C, a method for profiling of interactions around a so-called viewpoint, a region of interest. Their multiplexed 4C (MC-4C) method starts with a conventional 4C approach: the genome is cross-linked and digested, and fragment ends are ligated, but instead of cutting the resulting long DNA concatemers down to short fragments and amplifying those that interact with a viewpoint, the researchers kept the ligated DNA strands 2–5 kilobases long. The team used primers specific for the β-globin locus to amplify all fragments that interacted with this viewpoint, then sequenced the long reads on a MinION nanopore sequencer.

Conceptually the approach is simple, but de Ridder acknowledges that the computational analysis of multiple interactions "has given us a bit of a headache. The genome is a big place and multiway interactions become very sparse."



Multiway contacts along a mammalian allele and high-resolution Hi-C in yeast. Credit: Marina Corral Spence/Springer Nature.

De Ridder recalls that they started with long reads from a Pacific Biosciences machine, but the throughput was an order of magnitude too low. With the higher throughput of nanopore reads, he says, "we started to appreciate contacts that occurred at a much lower rate." To further boost the detection rate, they incorporated a few tricks, such as elimination of close-range interactions. "You can now identify the microtopology on the same allele. Computationally we can disentangle PCR duplicates from two biological occurrences," explains de Ridder. The order and orientation of the interacting regions on the reads tell the researchers whether they come from the same allele or are hybrids from mixing of two alleles.

Taking a closer look at the β-globin locus and its multiway interactions, the team found evidence for both cooperative and competitive interaction between enhancers and the genes they control, depending on whether the genes are active in the same stage of development.

De Ridder and de Laat plan to continue their collaboration and expand the approach to Hi-C and to profile multiway interactions genome-wide. De Ridder sees higher sequencing throughput as the main requirement and will switch to the larger PromethION machine to get more reads.

Meanwhile, in France, Roman Koszul at the Institute Pasteur approached the challenge of resolution in Hi-C not with longer reads from the mouse genome but with fragments of uniform length in budding yeast. Koszul and co-first author Héloise Muller are long-standing members of the synthetic yeast project Sc2.0, which has the ultimate goal of synthesizing the entire yeast genome. A few years ago Koszul wanted to combine his interests in synthetic biology and genome structure. "We were struggling with the analysis of Hi-C data; while working on computational tools we wanted to also work directly on the chromosome to alleviate some of the limitations we identified at the time," he says.

Koszul and his collaborators selected a 144-kilobase region on chromosome IV and introduced restriction sites on one allele to generate even-length restriction fragments for a Hi-C experiment. This allowed them to differentiate between the modified and the wild-type allele and to look at their structure during prophase 1 of meiosis, particularly at recombination hot spots. The experiment revealed the expected reorganization of all chromosomes into arrays of chromatin loops, and the engineered region allowed them to look specifically at homolog pairing. To their surprise, they saw fewer contacts between the homologous chromosomes than expected.

Koszul now plans to engineer the homologous region on the other chromosome IV with different restriction sites in order to follow both copies at high resolution. "We can then see the interplay between chromosome folding and recombination," he predicts.

He is also convinced that the approach is not limited to yeast, but could be applied to other organisms. "Not for full genomes, of course, or full chromosomes," Koszul says, "but for specific regions. It could be quite interesting to investigate repeat elements in a more complex genome."

From yeast to mammalian cells, the enigma of genome structure is unfolding. ❐

Nicole Rusk

Published online: 31 August 2018
https://doi.org/10.1038/s41592-018-0130-z

Research papers

Allahyar, A. et al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat. Genet.* **50**, 1151–1160 (2018).
Muller, H. et al. Characterizing meiotic chromosome structure and pairing using a designer sequence optimized for Hi-C. *Mol. Syst. Biol.* **14**, e8293 (2018).