GENOMICS

RNAs that shape the genome

Chromatin-associated RNA sequencing elucidates the role of RNAs in genome regulation.

Start local, then go global: this is the trajectory followed by Aaron Straight and his collaborators at Stanford University. "We had been interested in the idea that RNA might be structuring chromatin in some way," recalls Straight. In 2017, his team discovered that RNAs encoded by repeats remain associated in cis with the mitotic chromosome and recruit methyltransferases that help generate heterochromatin. This finding prompted a more general search for RNAs involved in genome regulation. The first requirement was a method to profile all genome-associated RNAs.

What initially started as a side project was boosted by a seed grant from the Stanford Center for Systems Biology to students and postdoctoral fellows in support of projects deemed risky. Jason Bell, a postdoc in the Straight lab, teamed up with David Jukam from Jan Skotheim's lab; Viviana Risca, working with Will Greenleaf; and three rotation students in the Straight lab—Owen Smith, Nikki Teran and Whitney Johnson. Together they developed an approach to enrich RNAs bound to the genome.

Their chromatin-associated RNA-sequencing (ChAR-seq) approach tackles the challenging aspects of generating a covalent linkage between DNA and RNA. The team first cross-linked RNA and chromatin in situ, then appended an adenylated and biotinylated linker to the 3' end of RNA, using an RNA ligase that will link only an adenylated DNA substrate, to avoid spurious RNA–DNA ligation. Then, after reverse transcription and digestion of the genomic DNA, they ligated the other end of the linker to DNA. Biotin allowed easy purification of the complex followed by amplification and sequencing. This procedure not only kept RNA and DNA together but also retained the polarity of the chimeric reads; the read sequence upstream of the linker came from RNA, the downstream sequence from DNA. The researchers could then map each segment to the transcriptome or genome and derive contact maps for cis or trans binding RNA.

To ensure that ChAR-seq worked as expected, the researchers first applied it to the fly genome and verified that they could detect noncoding RNAs known to coat the X chromosome in male flies.

Unexpectedly, they also found RNAs that bound every single chromosome. "At first, we thought it was an artifact," remembers Straight, but upon closer inspection they identified some of the RNAs as components of the spliceosome and realized they were looking at cotranscriptional splicing. The high resolution of ChAR-seq allowed the team to map the spliceosomes to the open reading frames they were associating with. Other ubiquitously binding RNAs remain to be characterized. "There are interesting classes of RNA that coat every chromosome; for some of them we have no idea what they are doing," says Straight. "There is a lot of biology here to dig into."

Another fertile area for ChAR-seq is the functional characterization of noncoding RNAs. Even if the function of a particular RNA is unknown, its binding site can be correlated with other marks in that region, such as histone modifications, that may give a clue as to the RNA's function. Straight is particularly excited about this capability, which he wants to apply to human genomes. "Of the roughly 20,000 noncoding RNAs, we know the function of about 50 of them," he says; "ChAR-seq is a great hypothesis generator."

Straight's team is also planning to use ChAR-seq in *Xenopus* embryos to map the RNAs that bind the genome in the absence of transcription and characterize the change in RNA occupancy as the genome is activated and transcription begins.

ChAR-seq joins similar approaches such as GRID-seq, developed in Xiang-Dong Fu's lab at UCSD, each with its unique strengths, each contributing to understanding of the interplay between RNA and chromatin to ensure proper genome function.

Nicole Rusk

Published online: 31 May 2018
https://doi.org/10.1038/s41592-018-0030-2

Research papers
Bell, J. C. et al. Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. *eLife* 7, e27024 (2018).