GENOMICS

Noncoding RNA's genomic hangouts

Long noncoding RNA interactions with chromatin can be mapped genome-wide using biotinylated tiling oligos.

The advent of tiling microarrays a decade ago liberated a profusion of noncoding RNAs from their previously shadowy existence in the genome. Early users of tiling arrays like Howard Chang at Stanford University became fascinated with these noncoding transcripts, which—despite their negative moniker—have been shown to actively regulate gene expression by several means. Many long noncoding RNAs (lncRNAs; typically over 200 base pairs) have been recently implicated in the epigenetic regulation of chromatin, in normal development and disease.

Whereas existing tools were focused on testing local lncRNA interactions, Chang wanted a method that would reveal lncRNA occupancy across the genome. "We were inspired by chromatin immunoprecipitation–sequencing as a powerful method for genome-wide protein localization and realized that an equivalent was needed for lncRNAs," he recalls. His group began with a simple premise. An oligonucleotide probe complementary to a lncRNA could selectively capture the neighborhood of lncRNA-associated DNA and protein using magnetic strepatividin beads that bind to a biotin moiety on the probe. Despite using highly stable Morpholino oligonucleotide probes, their first efforts failed. It was only by analyzing the capture efficiency of different probes targeting the same RNA that graduate student Ci Chu made a critical realization: lncRNAs are fragmented during the sonication step to make the chromatin soluble, such that a single probe can only pull down one fragment.

This eureka moment started the group thinking about their experience designing tiling arrays. This time, they tiled a lncRNA, devising complementary 20-nucleotide overlapping probes that avoided conserved regions. Designing a single probe requires detailed knowledge of secondary structure that can interfere with hybridization, but targeting the entire transcript circumvents this problem. They named their approach chromatin isolation by RNA purification and sequencing (ChIRP-seq). The method can be used to isolate DNA or protein, but to avoid isolating DNA from direct RNA probe-DNA hybridization, single-stranded RNA is gently removed by enzymatic digestion.

A potential drawback of using many short probes specific to a single lncRNA is the capture of similar off-target sequences. To get around this, the team reasoned that restricting analysis to sequences enriched by two independent probe sets would substantially reduce noise. In the resulting 'split-probe' strategy they only consider the intersection of data from non-overlapping sets of alternating 'even' or 'odd' probes that are assayed in separate experiments. "It is analogous to pulling down proteins using tandem affinity purification or multiple polyclonal antibodies," says Chang. "Corroboration by more than one pulldown gives higher stringency."

The strategy has paid off. The team tested their method on roX2, a ~600 nucleotide lncRNA that is known to associate with the male-specific lethal (MSL) protein complex in fruit flies. MSL binds many sites along the male X chromosome to upregulate gene expression in compensation for having one less copy of these genes in males than in females. High-resolution data from chromatin immunoprecipitation–sequencing of the MSL3 protein showed excellent correlation with roX2 ChIRP-seq data. Approximately 90% of binding sites were detected using both methods, and ChIRP-seq data were less noisy and had fine resolution, in the tens to hundreds of base pairs. Notably, ChIRP-seq data did not map to any spurious site among the autosomes, which comprise 80% of the genome.

ChIRP-seq has already contributed to our understanding of lncRNA biology. The group used the technique to map occupancy of TERC and HOTAIR lncRNAs with very interesting results. "We found that lncRNAs associate with numerous focal loci of several hundred base pairs, they are sequence-specific and at the informational level they underlie a lot of biological specificity," says Chang. These are still early days in the study of lncRNA interactions with chromatin, and many questions remain. ChIRP-seq promises to be a powerful tool to study how these regulatory transcripts, once ignored, assert their place in the genome.

Tal Nawy

RESEARCH PAPERS

Chu, C. *et al*. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* advance online publication (29 September 2011).

© 2011 Nature America, Inc. All rights reserved. npg