Protocol Report: Chromatin Immunoprecipitation Sequencing (ChIP-Seq)

Dieu My Nguyen

MCDB 5521 — February 28, 2018

ChIP-Seq is a sequencing method that combines chromatin immunoprecipitation assays with parallel DNA sequencing to identify genome-wide binding sites of specific transcription factors and proteins that interact with DNA. This method contributes to an epigenetic perspective of biological processes and disease states. I would like to study this sequencing method and its data in exploring the effects of intruder response on brain chromatin accessibility in the honey bee (*Apis*).

The general ChIP-Seq protocol is:

1. Cross-link and cell harvest:
Cross-link proteins to their bound DNA by formaldehyde treatment, *in vivo*. This step is time-dependent and requires optimization, to avoid excessive cross-linking which reduces antigen accessibility and sonication efficiency for later steps. The cross-linking reaction is terminated by adding glycine to quench the formaldehyde.

2. Sonicate:
Shear DNA strands by sonicating lysate. This step creates an average fragment size of 200-1000bp.

3. Determine DNA concentration and fragment size:
Use the sonicated chromatin samples to calculate DNA concentration and fragment size. Samples are treated with RNase A and proteinase K.

4. Immunoprecipitate:
Add antibodies bound to magnetic beads to immunoprecipitate target protein and precipitate DNA.

5. Purify DNA and reverse cross-links:
Disrupting the cross-links between the DNA and the proteins help with DNA purification, which is done by using a PCR purification kit or phenol:chloroform extraction.

6. Sequencing:
Add oligonucleotide adaptors to small regions of DNA that are bound to the proteins of interest. Sequence the immunoprecipitated DNA to analyze for DNA binding sites. Map sequence to genome.

Advantages of ChIP-Seq:
- High base-pair resolution of protein-binding sites
- Can map specific transcription factors and proteins
- No contamination by unbound DNA by using exonuclease

Disadvantages of ChIP-Seq:
- DNA-protein complexes of interest can be diluted by nonspecific antibodies
- Must know protein of interest

Like other high-throughput sequencing methods, ChIP-Seq outputs large datasets that require adequate and appropriate analysis. Some analysis methods have been developed for ChIP-Seq outputs: peaking calling methods for predicting DNA-binding sites; differential peak calling for predicting differences between ChIP-Seq signals and peaks using Hidden Markov Models; and Model-based Analysis for ChIP-Seq (MACS) for modeling the shifts of ChIP-Seq tags and increasing spatial resolution of predicted binding sites.

An example of data generated by ChIP-Seq:

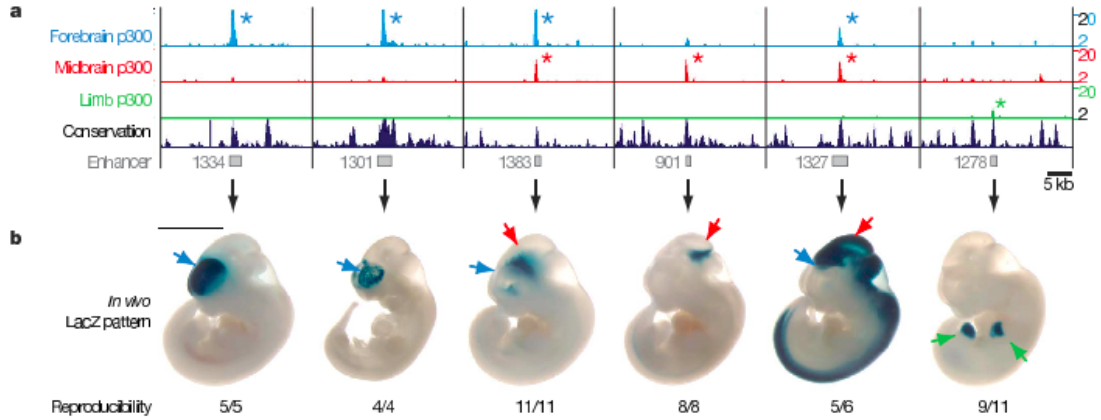

Figure 1: ChIP-Seq accruately predicts *in vivo* enhancers by p300 binding in embryonic tissues [1]

References

[1] Visel. Axel et al. "ChIP-seq accurately predicts tissue-specific activity of enhancers". In: *Nature* 457.7231 (2009), pp. 854–858.

[2] Abcam. *ChIP protocol for cross-linking chromatin immunoprecipitation (X-ChIP)*.

[3] Bailey Timothy L and Machanick Philip. "Inferring direct DNA binding from ChIP-seq". In: *Nucleic Acids Research* 40.17 (2012), e128. DOI: 10.1093/nar/gks433.

[4] Blow Matthew J et al. "ChIP-Seq identification of weakly conserved heart enhancers". In: *Nature Genetics* 42.9 (2010), pp. 806–810. DOI: 10.1038/ng.650.

[5] Illumina. *DNA sequencing methods collection*.

[6] Shi Yuan-Yuan et al. "DNA methylation comparison between 4-day-old queen and worker larvae of honey bee". In: *Journal of Asia-Pacific Entomology* 20.1 (2017), 299—303. DOI: 10.1016/j.aspen.2017.01.014.

[7] Shpigler H.Y et al. "Behavioral, transcriptomic and epigenetic responses to social challenge in honey bees". In: *Genes, Brain and Behavior* 16 (2017), pp. 579–591. DOI: 10.1111/gbb.12379.