

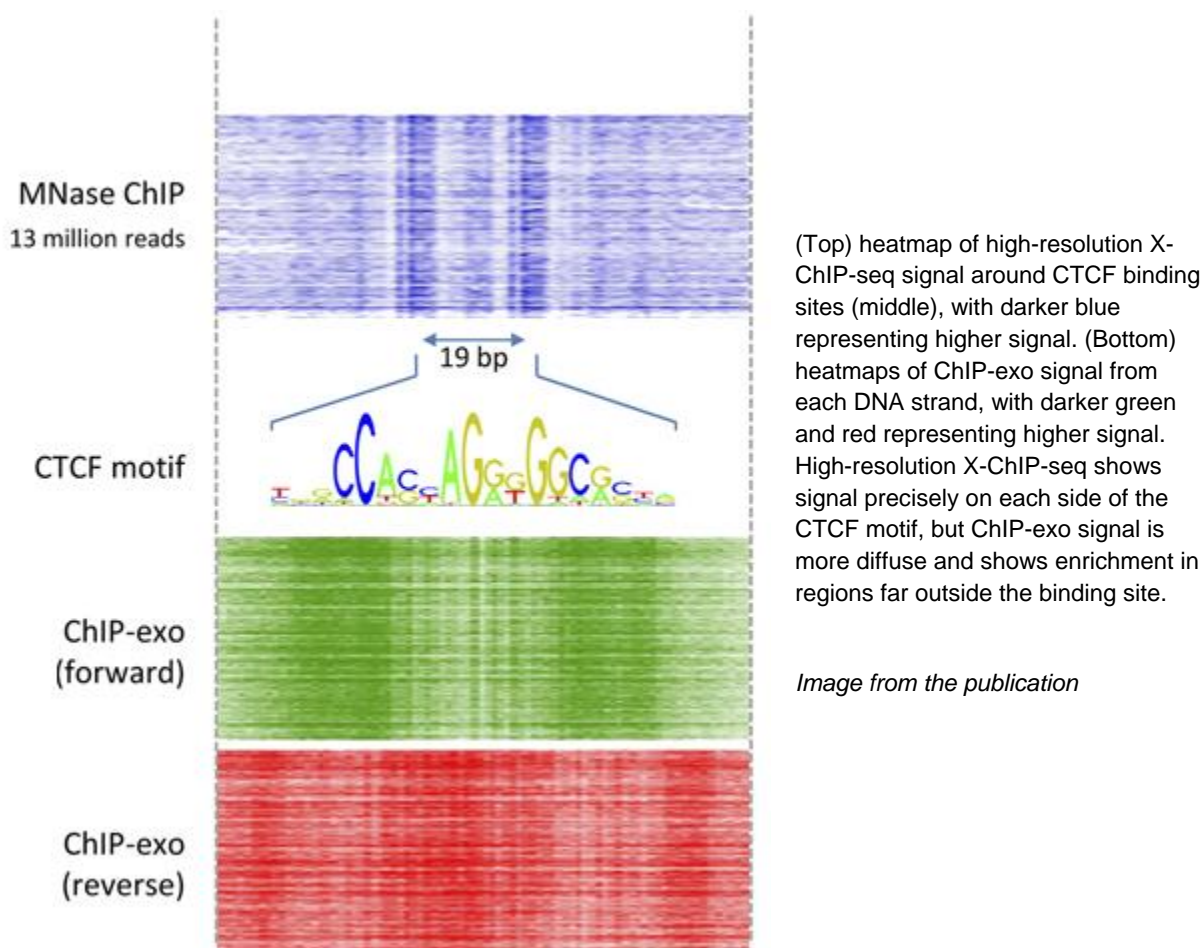
August 17, 2015

SCIENCE SPOTLIGHT

# ChIPping away at protein-DNA interactions in high definition

Aug. 17, 2015

GE Zentner



Interaction of various proteins with the genome is critical for DNA-based processes including replication, transcription, and repair. Thus, defining specific sites of protein-DNA interaction is of great importance for understanding these fundamental cellular processes, as well as their dysregulation in disease. The most common technique used to map genome-wide protein-DNA interactions is chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). ChIP-seq has been used in thousands of studies to map protein binding sites, but suffers from low resolution due to the use of sonication to fragment chromatin. To improve the resolution of ChIP-seq, a previous study described the addition of exonuclease digestion steps to remove flanking

DNA not protected by proteins. This modification of ChIP-seq, termed ChIP-exo (Rhee and Pugh 2011) greatly improves the resolution of binding sites, but are technically challenging and require complicated computational analysis. Dr. Peter Skene, a postdoctoral fellow in the lab of Dr. Steven Henikoff (Basic Sciences Division), previously used a simple modification of conventional ChIP-seq, termed high-resolution X-ChIP-seq, in which micrococcal nuclease (MNase) digestion is used to digest away DNA not protected by protein. Dr. Skene previously used this approach to generate high-resolution maps of RNA polymerase II (RNAPII) binding to the mouse genome. In a recently published eLife Research Advance, Dr. Skene compared the performance of high-resolution X-ChIP-seq to other ChIP-seq methodologies, finding that it compares favorably with both standard ChIP-seq and ChIP-exo.

To evaluate the performance of high-resolution X-ChIP-seq in mapping RNA polymerase II, Dr. Skene profiled RNAPII binding in *Drosophila melanogaster* cells and compared the output to data generated by either conventional ChIP-seq or 3'NT, a high-resolution RNA mapping technique that gives the exact base pair (bp) position of the last nucleotide added to a growing mRNA molecule by RNAPII. High-resolution X-ChIP-seq and 3'NT were in excellent agreement, showing RNAPII binding ~35 bp upstream of the transcription start site (TSS), consistent with evidence suggesting that transcriptionally engaged RNAPII stalls just downstream of the TSS. In contrast, conventional ChIP-seq showed prominent RNAPII enrichment directly over the TSS. These contrasting results highlight a drawback of the use of sonication for chromatin fragmentation: regions of the genome that are depleted of nucleosomes, including gene promoters and TSSs, are more susceptible to breakage by sonication and are thus overrepresented in conventional ChIP-seq datasets.

The length of DNA protected by RNAPII and other factors is relatively short. Such small DNA fragments are difficult to retain during sequencing library preparation and so constitute a relatively small fraction of total sequenced reads, necessitating greater sequencing depth for good representation of these fragments of interest. Dr. Skene therefore introduced a method to selectively enrich for small fragments during library preparation that yielded a ~25-fold enrichment of these short fragments.

To test the performance of this short fragment purification scheme, he performed high-resolution X-ChIP-seq for the CCCTC-binding factor (CTCF), a well-characterized transcription factor. Comparison of small fragments from high-resolution X-ChIP-seq to conventional ChIP-seq data yielded ~4-fold narrower peaks in the size-selected data, consistent with high-resolution X-ChIP-seq significantly improving binding site resolution. Comparison of CTCF high-resolution X-ChIP-seq and ChIP-exo showed that, even with exonuclease digestion, ChIP-exo gives slightly lower

resolution of CTCF binding sites that high-resolution X-ChIP-seq. This may be because sonication produces fragments 200-400 base pairs in length that may have other proteins crosslinked to them, which in turn serve as barriers to exonuclease digestion. In contrast, MNase is an endo/exonuclease and can thus cut DNA between crosslinked proteins, providing even greater resolution than ChIP-exo.

"ChIP-exo provides much higher resolution than conventional ChIP-seq, but is very complicated," said Dr. Skene. "Conceptually, X-ChIP-seq is no more difficult than conventional ChIP-seq while still providing very high resolution, so any lab familiar with ChIP-seq can easily put X-ChIP-seq into practice."

[Skene PJ, Henikoff S](#). 2015. A simple method for generating high-resolution maps of genome-wide protein binding. *eLife* 4:e09225.

See also: [Rhee HS, Pugh BF](#). 2011. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* 147(6):1408-1419.

This work was supported by funding from the Damon Runyon Cancer Research Foundation and the Howard Hughes Medical Institute.