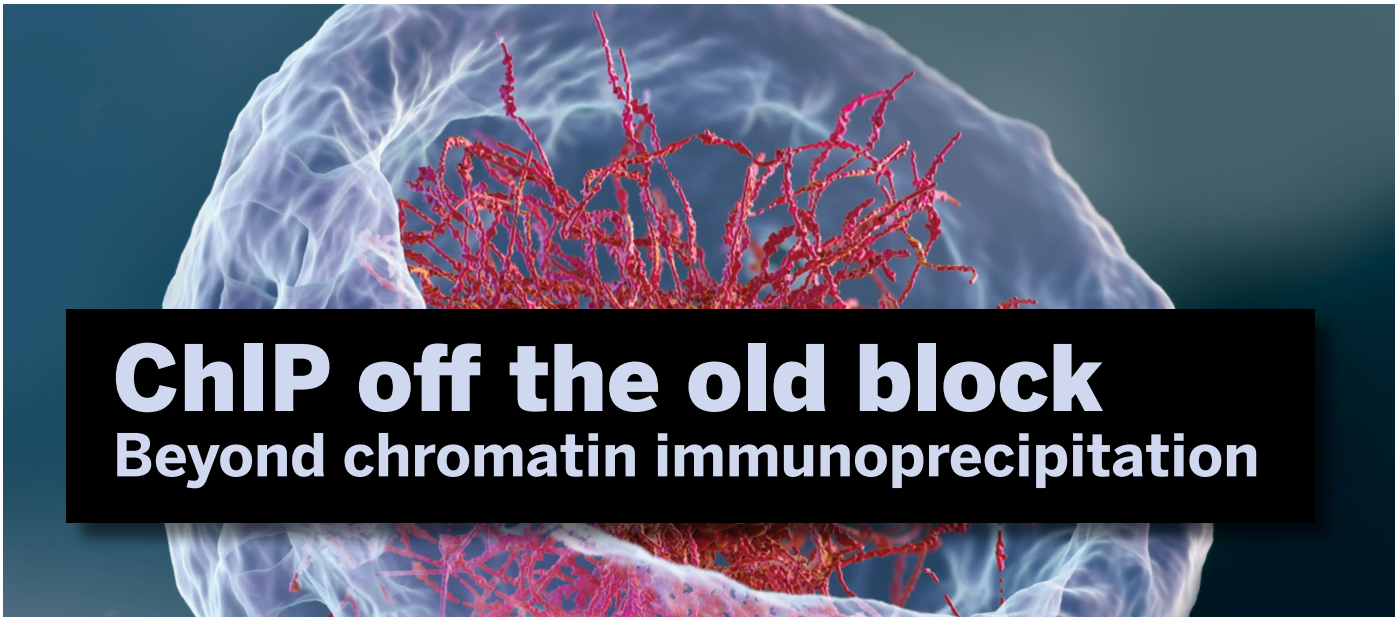

ChIP off the old block
Beyond chromatin immunoprecipitation

A host of techniques are building from a classic method—chromatin immunoprecipitation (ChIP)—to assess what binds to DNA and where. Emerging techniques whittle down the size of samples, interrogate DNA-bound protein complexes, or more closely assess the nucleotides involved. All of these approaches aim to overcome long-standing limitations of ChIP, and broaden the questions scientists can ask about gene regulation, development, and disease. **By Charlotte Schubert**

Chromatin immunoprecipitation, one of the most widely used techniques in molecular biology, was invented over 30 years ago—and some things about it have changed, while others have stayed the same.

The basic protocol is still similar to one developed in the 1980s, involving crosslinking proteins to DNA with formaldehyde and then fragmenting the DNA. A DNA-interacting protein is immunoprecipitated using an antibody, the crosslinks reversed with heat, and the associated DNA analyzed. Researchers later linked the technique to deep sequencing, developing the ChIP-seq technique to probe protein-DNA interactions at the genomic scale.

ChIP has been harnessed to address how transcription factors operate, how histones modulate gene expression, and other basic questions with implications for biological development and disease. In September, C. David Allis and Michael Grunstein won the prestigious Lasker award for their work on histones—research that relied on ChIP. And ChIP-seq is now a cornerstone of the ENCODE (ENCyclopedia Of DNA Elements) project, an effort to map regulatory regions of the genome in various cell types.

Chromatin biologists are also developing an array of spin-off or parallel technologies to go beyond what ChIP and ChIP-seq offer—to examine complexes of proteins, to more accurately assess the exact nucleotides a factor binds to, to look at small pools of cells, and to begin, tentatively, to assess protein-DNA interactions at the single-cell level.

All of these techniques aim to do things that ChIP-seq alone cannot, or does only sluggishly. And all of them have the same basic goal: to find out what molecules are associated with DNA and where.

"We really don't understand the fundamental principles by which regulatory functional sequences in our genome determine where and when genes come on," says Bradley Bernstein, director of the **Broad Institute's Epigenomics Program** in Cambridge, Massachusetts. He adds that ChIP is "limited in many ways. And so there are these efforts to try and innovate new approaches or adapt the technology in new ways."

Making it work

"Calling [ChIP-seq] a dark art is too much," says Nir Friedman, a professor of computer science and biology at **The Hebrew University of Jerusalem,** Israel. But despite it being a commonly used technique, "very few people are patient enough to calibrate their experiments," he says.

ChIP-seq experiments generate a lot of noise, notes Friedman. Formaldehyde can crosslink uninvolved molecules, antibodies can pull down nontarget proteins, and sonication—the most common way to break up DNA—tends to break up DNA that is in an open conformation. Says Friedman, "It can be that more than half of what you end up sequencing or looking at is nonspecific binding."

ENCODE publishes guidelines for assessing the quality of antibodies and screening out meaningless data, Friedman notes. The project also provides access to recent computational tools that are used, for instance, to normalize data to controls and to identify "peaks" or regions of possible DNA binding.

Choosing the right antibody, in particular, can be challenging, notes Michael-Christopher Keogh, chief scientific officer at **EpiCypher,** an epigenomics company in Research Triangle Park, Durham, North Carolina. Keogh was involved in a recent study showing that many antibodies popular for histone research perform poorly in ChIP, for instance, binding to off-target epitopes. The study also proposes validation steps beyond the ENCODE guidelines.

Some researchers bypass the antibody problem by engineering an epitope tag onto their target, as with CETCh-seq (CRISPR epitope tagging ChIP-seq). A long-standing technique, DamID (DNA adenine methyltransferase identification), involves engineering factors to tag neighboring DNA with molecular marks. cont.>

Upcoming features
Next-Gen Sequencing: RNA Seq—February 22 ■ Microscopy: Multiphoton Microscopy—March 22 ■ Molecular Biology: Epitranscriptomics—May 17

Other researchers are still improving on the basic ChIP technique, such as Alon Goren's group at the University of California San Diego's Department of Medicine, which has fully automated ChIP-seq. Goren says that the optimal antibody concentration can vary substantially among antibodies and cell types, and that some steps, such as reverse crosslinking, are unnecessary. The **Goren Lab** has also shown that monoclonal antibodies have the edge over polyclonal antibodies.

But even when fully optimized, ChIP-seq is still fundamentally limited. For instance, it generally requires 100,000 or more cells to assess transcription factors and 10,000 or more to assess histone proteins, says Goren. And in most cases, the method is geared to look at one protein, one antibody at a time.

Says Goren, "Once we are able to shift the view from thinking about proteins to thinking about complexes, we will get a much better understanding of how biology works and what happens in disease."

Getting a handle on protein complexes

One approach to examine complexes is to combine ChIP-seq with mass spectrometry (MS), using methods such as RIME (Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins) and ChIP-MS.

One drawback of these methods, however, is that proteins not associated with DNA can also be pulled down by immunoprecipitation, says David Steger, a molecular biologist at the **University of Pennsylvania,** in Philadelphia. Instead, says Steger, "What everybody is trying to develop is locus-specific proteomics at a particular enhancer."

One emerging technique to assess chromatin-bound complexes is ChIP-SICAP (selective isolation of chromatin-associated proteins), developed by Jeroen Krijgsveld's team at the **German Cancer Research Center** in Heidelberg, Germany, and his colleagues. The technique involves tagging antibody-bound DNA with biotin, which is then pulled down with biotin-binding streptavidin beads before mass spec.

Steger is applying ChIP-SICAP to examine proteins bound to enhancers that drive the transition of mesenchymal stem cells to adipocytes. Says Steger, "What we are trying to do is identify an enhancer proteome."

Other approaches harness the gene-editing system involving Cas9, which recognizes guide RNAs targeted to specific DNA sequences. Researchers have tagged Cas9 with biotin or an enzyme that promotes the biotin labeling of nearby proteins, which are analyzed by MS. This approach has also been deployed to reveal interactions between distant genomic elements. Such methods could potentially extend the repertoire of ChIP-seq–related methods that can assess the 3D architecture of the genome, such as Hi-C, ChIA-PET (Chromatin Interaction Analysis by Paired-End Tag sequencing), and HiChIP, and more recent methods such SPRITE (Split-Pool Recognition of Interactions by Tag Extension).

These emerging methods to assess DNA-bound complexes promise to sharpen biologists' view of gene regulation. But when harnessed to MS, they face the limitation that MS does not readily detect low-abundance proteins, notes Steger. Moreover, not all newer techniques are accessible to the nonexpert.

Several companies offer services for outsourcing of more-established techniques such as RIME or ChIP-seq. Companies that offer ChiP-seq and related services include **Active Motif,** in Carlsbad, California; **Diagenode** in Liege, Belgium and Denville, New Jersey; and Beijing-based **Novogene.** These and other companies also offer ChIP-seq kits and components, though many labs generate their own reagents. Keogh also notes that in-house research can mean greater control of optimization steps.

IMAGE: © IMAGENTLE/SHUTTERSTOCK.COM

Care with experimental parameters during ChIP experiments can itself boost the quality of data, says Cigall Kadoch, whose group studies several large macromolecular protein complexes at **Dana-Farber Cancer Institute** and **Harvard Medical School** in Boston, Massachusetts.

Kadoch is careful to optimize the concentration of formaldehyde used to crosslink proteins together with each other and DNA, for each antibody and cell type. When choosing antibodies, she notes whether the epitope is predicted to be accessible on the surface and therefore amenable to immunoprecipitation. And to see the DNA footprint of a fully assembled complex, she advises choosing an antibody to a protein that is added on late in the assembly process. "These are the things that make or break a project," Kadoch says.

Zeroing in on factor binding

Another technique Steger uses is ChIP-exo (ChIP exonuclease). He deploys it to identify—at base-pair resolution—where various factors bind the genome.

This technique starts with DNA fragmentation by sonication. An exonuclease chews up the DNA (in the 5'–3' direction) to the edge of where the DNA is linked by formaldehyde to its bound protein. This approach results in a sharp DNA "footprint" for bound factors, which can be more exact than the inferred motifs generated computationally using ChIP-seq.

"We always start with ChIP-seq, and as our questions evolve, we move to ChIP-exo," says Steger, who has used the technique to assess the binding of the glucocorticoid receptor to DNA. The receptor binds as a dimer to two abutting, short DNA sequences. Steger was able to resolve binding of one monomer, which he was unable to do with ChIP-seq.

The team of Frank Pugh, a professor of biochemistry and molecular biology at Penn State University in University Park, Pennsylvania, recently simplified its ChIP-exo method and adapted it to the commonly used **Illumina** sequencing platform. Similarly, Julia Zeitlinger, associate investigator at Stowers Institute for Medical Research in Kansas City, Missouri, and her colleagues, have published a related technique, ChIP-nexus. Both developments are "promising," says Michael Snyder, chair of genetics and director of the **Stanford Center for Genomics and Personalized Medicine** at Stanford University, in California. cont.>

Featured participants

Active Motif
www.activemotif.com

Altius Institute for Biomedical Sciences
www.altius.org

Benner Lab, University of California, San Diego
homer.ucsd.edu/BennerLab

Broad Institute Epigenetics Program
www.broadinstitute.org/epigenomics

Dana-Farber Cancer Institute
www.dana-farber.org

Diagenode
www.diagenode.com/en

EpiCypher
www.epicypher.com

Fred Hutchinson Cancer Research Center
www.fredhutch.org/en.html

German Cancer Research Center
www.dkfz.de/en/index.html

Goren Lab, University of California, San Diego
goren-lab.github.io/index.html

Harvard Medical School
hms.harvard.edu

Illumina
www.illumina.com

Novogene
en.novogene.com

Stanford Center for Genomics and Personalized Medicine
med.stanford.edu/scgpm.html

The Hebrew University of Jerusalem
new.huji.ac.il/en

University of Pennsylvania
www.upenn.edu

Additional resources

ENCODE
www.encodeproject.org/pipelines

Going smaller and deeper

Researchers have been able to shave off the number of cells required for ChIP-seq by adjusting experimental parameters, such as using a high-quality antibody, says Friedman.

Some techniques, including one developed by Friedman, use barcoded sequencing adaptors, enabling decreased sample size. And a new technique called "ChIPmentation" reduces the steps involved in making sequencing libraries.

One method making its way to new labs is CUT&RUN (Cleavage Under Targets and Release Using Nuclease), developed by Steve Henikoff and his colleagues at **Fred Hutchinson Cancer Research Center** in Seattle, Washington. This method dispenses with crosslinking by formaldehyde as well as DNA shearing with sonication. Instead, an antibody against the target is tethered to micrococcal nuclease (MNase), which is activated by calcium to cleave the DNA on either side of the target. The resulting DNA fragments are sequenced.

"You get a large signal-to-noise reduction" as compared to ChIP-seq, says John Stamatoyannopoulos, director of the **Altius Institute for Biomedical Sciences** in Seattle. That's in part because of the clean cutting of DNA by the nuclease, with low levels of off-site cutting. As a result, CUT&RUN typically requires fewer DNA sequencing reads than ChIPseq—reducing costs—and can be applied to much lower cell numbers. Henikoff's team recently applied the technique to 1,000 cells for a transcription factor and 100 cells for a histone modification. Stamatoyannopoulos says the method has largely supplanted ChIP-seq in his labs.

CUT&RUN also has advantages beyond low cell numbers. Stuart Orkin, a molecular biologist and professor at Harvard University, touts the technique for its "essentially nucleotide-level" resolution. With minor tweaks to computational tools, his group was able to differentiate the closely spaced binding sites of a transcription factor involved in controlling the expression of fetal hemoglobin. Prior to obtaining this result, he was unable to immunoprecipitate the transcription factor with conventional ChIP, possibly because formaldehyde crosslinking hid the epitopes. CUT&RUN "worked right off the bat," he says.

Henikoff's lab has adapted the technique to assess long-range, 3D DNA interactions, and to perform immunoprecipitation on the cleaved-out fragments using a second antibody. The group interrogated two molecular features on the same protein complex—an approach that could help resolve questions such as which combinations of histone marks are associated with various gene states.

Working at the single-cell level

For many molecular biologists, including chromatin researchers, the single cell is the final frontier.

"We need to come up with precise, deterministic ways of directly evaluating single-molecule interactions systematically in single cells," says Bernstein. "It is a long-term goal." Single-cell data could potentially track DNA-binding factors as cells exit the stem-cell state during development, or in tumors with high levels of cellular heterogeneity.

Several approaches are edging closer to this goal. However, "A lot of the single-cell methods have limited sensitivity," says Christopher Benner, a genome biologist at the **University of California, San Diego.** Bernstein and his colleagues for instance, generated single-cell ChIPseq data using a microfluidic system and barcoding. From each cell, the technique captured between 500–10,000 unique "reads," representing a DNA-binding event, in contrast to the millions of reads captured with populations of cells.

Several methods can also yield data on the active regions of the genome. ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) deploys a hyperactive transposase that integrates into the genome in open chromatin regions and introduces sequencing adaptors. ATAC-seq can be adapted to single cells and the resulting sequence data interrogated with a variety of computational approaches. These approaches can, for instance, identify promoter sequences—or identify patterns in experiments simultaneously knocking out DNA control elements or assessing RNA expression.

ATAC-seq suffers from drawbacks, such as incomplete integration into accessible regions, notes Snyder. But the technique can be powerful: It can be deployed to insert and image fluorophores, resulting in imaging data on 3D genomic organization prior to sequencing, notes Snyder.

Snyder points to imaging work in labs such as that of Alistair Boettiger at Stanford, who is developing ways to simultaneously image genetic elements and nascent RNA transcripts with super-resolution imaging, to assess which elements promote or quell gene expression. "Imaging is the future," adds Stamatoyannopoulos. Other researchers note that as "third-generation" nanopore sequencing improves, ChIP-like methods will be developed to plug into it.

"We may need entirely orthogonal ways of doing this," says Bernstein of single-cell chromatin analysis. That goal will likely will be accomplished successfully in the end, he says, with "technologies that are a radical departure from what we are using now."

Charlotte Schubert is a freelance journalist based in Seattle. From 2015–2016 she also worked at the bench in Steve Henikoff's lab, on a project unrelated to the work highlighted here.