

What to do about those immunoprecipitation blues

ChIP-seq, DIP-seq and related techniques are informative genome-wide assays, but they don't always work as planned.

Vivien Marx

When the chromatin immunoprecipitation (ChIP) or DNA immunoprecipitation (DIP) blues hit, it's good to realize you're not alone. ChIP is part of the winding path of characterizing gene function. To find where transcription factors (TFs) or histone marks bind to genomic loci of interest, labs might choose ChIP-seq, which involves cross-linking, then shearing chromatin, immunoprecipitation with an antibody that recognizes a TF or histone mark of interest, followed by sequencing. DIP plus sequencing (DIP-seq) can be used to locate DNA changes such as methylation. ChIP-seq and DIP-seq have been the “cornerstone of epigenetics research” since the field moved from gene-specific epigenetics to the genome-wide approaches of epigenomics, says Colm Nestor, a researcher at Linköping University. Both techniques are cheap, relatively easy to set up, and widely used “because they work very well—when well controlled, that is,” he says.

With ChIP-seq, antibodies are sometimes not as specific as a lab might have first assumed¹. Some genomic loci can be ‘sticky’ in too many wrong ways, or masked protein epitopes are not ‘sticky’ when they should be. As Dana Farber Cancer Institute researchers Xiaole Shirley Liu and Clifford Meyer point out, some of ChIP-seq's lurking bias issues are a lack of antibody specificity, issues with chromatin shearing, cross-linking and sequencing². To battle ChIP-seq and DIP-seq artifacts, experimental troubleshooting is unavoidable. Challenges are not always the antibody's fault, but when labs troubleshoot, antibodies need to be considered. Here are two cases of immunoprecipitation (IP) headache, and researchers at two antibody companies comment on ChIP-seq and DIP-seq challenges.

ChIP-seq data confidence rattled

ChIP-seq gives labs a crack at genome-wide analysis because “hundreds and thousands of binding sites are displayed in context,” says Peter Becker, a researcher at Ludwig Maximilians University in Munich. For example, the comparison of ChIP-seq profiles lets labs generate hypotheses about



Immunoprecipitation can be used to locate DNA changes such as methylation or places where transcription factors bind, and can involve the struggle with artifacts. Credit: A. Lentini, Linköping University

differential binding. His former PhD student Dhawal Jain, now a postdoctoral fellow in the lab of Harvard University researcher Peter Park, calls ChIP-seq experiments essential in functional genomics because of the insight they deliver on the biological roles of TFs. ChIP-seq is his first experiment of choice for characterizing any novel DNA-binding factor, and it's crucial to get data “with high confidence.” Both he and Becker got their data confidence rattled.

When Jain was in the Becker lab, the researchers studied nucleosome remodeling. Nucleosomes are complexes of DNA wound around histones, and nucleosome remodelers open transcription ‘windows’ at certain DNA locations. The remodelers only transiently interact with DNA, so they are tough to capture and map. Applying ChIP-seq with antibodies directed at parts of these remodeling factors, the team expected to see, and saw, broad peaks corresponding to the remodelers, says Jain. There were also

sharply defined peaks in open chromatin regions, where nucleosomes are usually lacking. The peaks “were striking to us,” he says, but not striking in a good way.

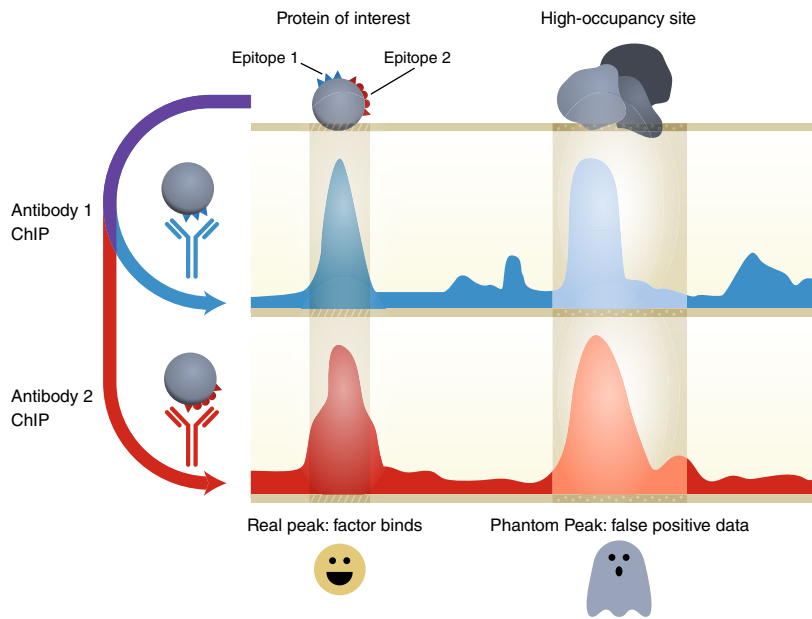
They investigated and identified around 3,000 loci that had non-specific ChIP enrichment³. To validate the experiment, they knocked out these two remodeling factors, and saw similar ChIP-seq profiles. They had found false positive signals: Phantom Peaks, which affect their data quality and, possibly, data in many labs and databases. Most Phantom Peak sites were at promoters, where many proteins bind with disordered domains, says Becker, and where non-specific antibody-binding is the likely culprit. These low ChIP peaks typically go unnoticed in experiments with strong, specific profiles, says Becker. “But as soon as your ChIP reaction is inefficient for any reason, and your specific peaks are low, Phantom Peaks show up.” The team advises interpreting ChIP-seq peaks with caution.

Their list of Phantom Peaks is part of their paper as a resource for others. To rule out Phantom Peaks, they recommend knockout or knockdown experiments. Sometimes these are not feasible if the factor under investigation is essential, says Jain. Input controls can be helpful, says Becker.

Even when labs have established antibody specificity through western blots (WB) and immunofluorescence (IF), non-specific binding can be a problem. The team ran into just such issues. “To add more to it, oftentimes, the antibody that works in one validation assay does not work in the other,” says Jain. Labs then try to validate in either WB or IF and try to reproduce the ChIP-seq profile with an orthogonal antibody. But the orthogonally generated ChIP-seq profiles do not necessarily remove Phantom Peak artifacts. They also found enrichment of a ChIP-seq signal with completely unrelated antibodies, such as anti-GFP. Overall, Jain says he sees a growing awareness about such ‘hyper-ChIPable’ regions of the genome that lead to false positive data. One issue with the nucleosome-remodeling factors they were studying, says Becker, notably ACF1 and RCF1, is that they do not cross-link well to chromatin given their transient interactions with DNA. Fundamentally, the antibodies in their experiment were fine, he says, but there was non-specific binding. Since that first work, they have tethered the protein ACF1 to DNA with a fused DNA-binding domain and could “ChIP it” just fine, he says. The issue was essentially noise related to non-specific antibody-binding.

DIP-seq peculiarities

Antibody-based techniques are useful for investigating newly discovered epigenetic changes such as N^6 -methyladenine, since it's not hard to raise antibodies against new modifications, says Linköping University's Nestor. “Consequently, researchers are spared the considerable hassle of having to develop modification-specific methods each time we detect a novel epigenetic mark in mammals.” He has been using DIP-seq for a decade, immunoprecipitating methylated DNA chunks and sequencing them. Nestor, his PhD student Antonio Lentini, and others on the team didn't set out to find a major DIP-seq flaw, but that's what happened. According to the team, between 50% and 99% of enriched regions in DIP-seq data might be false positives. They found non-specific enrichment of areas of the genome called short tandem repeats (STRs)⁴. Other known sources of bias include the fact that the 5-methylcytosine (5mC) antibody labs mainly use preferentially enriches genomic regions of low CG content and the most frequently used 5-hydroxymethylcytosine



ChIP-seq data can be littered with ‘Phantom Peaks’ that do not correspond to the immunoprecipitated protein of interest. The use of two orthogonal antibodies for pulldown can be a failsafe. But that does not always work. For example, the antibodies might bind non-specifically to proteins occupying the open chromatin. Credit: D. Jain, Harvard Medical School; P. Becker, LMU Munich; E. Dewalt/Springer Nature

(5hmC) antibody does the same at highly modified regions of the genome.

In their analysis of published DIP-seq studies, they saw that low-abundance genomic modifications, such as 5-formylcytosine and 5-carboxylcytosine, have the highest false positive rates. They had been looking at just those modifications to explore the role of DNA methylation in human T cells. They noticed “peculiar” binding patterns in their DIP immunoglobulin G (IgG) controls. As much as 95% of published work does not include an IgG control, they found. “As a researcher that has consistently used DIP-seq without an appropriate IgG control, I am in no position to point fingers at others,” says Nestor.

The team validated their results with embryonic stem cells lacking 5mC and 5hmC modifications. The antibodies still bound to STR sites. Perhaps single-strand DNA is binding to IgG antibodies. Lentini speculates that the STR binding may offer functional clues and that antibodies might be engineered to inhibit this binding. “Until then, our best option is to be aware of these issues and control for them,” he says. Avoiding false positives and correcting for this observed off-target binding, the team says, is best handled by normalizing DNA-modification enrichment to an IgG control, which increases the signal-to-noise ratio threefold, and labs can detect more subtle changes in DNA modifications. ‘Input’

has been the standard control, says Nestor, referring to the chromatin sample pre-IP. As a postdoctoral fellow, he remembers being advised to avoid IgG controls because they add too much noise to experiments. But the researchers say input is a highly inconsistent control and that the 5-modC landscape in mammalian genomes has been greatly overestimated by DIP-seq. Although Nestor and others use computational tools to filter the DIP-seq data, he is concerned about large errors creeping into datasets.

Another challenge is bacterial contamination. Cell cultures, often infected with *Mycoplasma* and *Escherichia coli*, show high levels of the modification N^6 -methyladenine (6mA), which is rare in mammals and frequent in microbes. 6mA “is still an enigma,” says Nestor. It's hard to detect, no one technique detects it reliably in mammals, and it remains a puzzle to be explored. He wonders “what the function of this incredibly shy member of the DNA modification club could be.” Andrew Chalmers, who is at the University of Bath, recommends that when scientists select antibodies for experiments, including ChIP-seq, they look at previous experiments with these reagents. Chalmers co-founded CiteAb, an online resource for finding published literature and company information about antibodies. The “five pillars paper” on antibody validation is not perfect, he says, but is a good framework for discussing and analyzing

performed validation⁵. Often, the planned experiment will not be among those previously done, but the available data and hard-earned lessons from others can be useful guides.

Knockout/knockdown

Experiments with antibodies blend on-target and off-target effects, says Velibor Savic, who leads the research and development teams at the antibody company Abcam. Experimentalists need to reduce off-target signals, also in ChIP, he says. ChIP-seq is sensitive, which can render off-target signals more visible than other techniques. That gives knockdowns and knockouts critical importance in ChIP-seq. “With this,” he says, “it is easier to tease out target from off-target signals and identify the correct positive peaks.”



Enrichment in the wrong places is one reason that he recommends using a highly validated antibody for ChIP experiments, says Chris Fry.

Credit: D. White

Even transient interactions can have a profound effect. “Such interaction may be poorly captured via ChIP,” says Savic. Results might blend into background because only some of the interaction sites are occupied at a given time. It can help to explore ways to stabilize the interactions. Using knockdown or functional knockout of the target would address both the off-target binding and non-specific interactions. “In most cases, this would nicely define the locations of the target protein within the experimental sample and the patterns of background in the absence of it,” he says. ChIP delivers a static picture of the average target-to-DNA interaction profile in a very dynamic system, he says. “As much as one can’t judge a movie from one still image, one can’t draw dynamic conclusions from a single ChIP experiment,” he says. The journey to functional insight only begins with ChIP-seq data, says Savic.

Cross-linking

Cross-linking, says Savic, is “one of the most critical variables” that need optimization in any ChIP program. Formaldehyde is the standard cross-linker, but as a short-range cross-linker it has limitations. It’s best suited for linking DNA and proteins and for localizing proteins that directly and strongly

interact with DNA, such as histones. But transient interactions, or those mediated through a protein complex, may not be well captured this way. More optimized, intermediate to long-range cross-linkers might be better. “Using these, either singly or in combination with formaldehyde, may be a critical step in boosting cross-linking efficacy to the level usable for subsequent steps,” he says. A protein can be cross-linked to the DNA it interacts with and there can be non-specific cross-linking with nearby DNA sequences, which would be randomized across the cell population and blend with the background signal, says Savic. But unexpected peaks in sequencing depend on the immunoprecipitated target. With non-random interactions such as those with sequence-specific promoters, a DNA segment can be ‘pulled down’ with the target protein even when it is not directly bound to it. Shearing and sonication quality are other important factors. Some structures will cross-link or sonicate better or worse, depending on chromatin condensation, compaction, protein–DNA interaction and many other factors, he says. These might end up in the IP reaction as fragments larger than 1 kb of DNA with an elevated chance of off-target pulldown. And “sonication-resistant” regions can deliver false positive results.

ChIP protocols differ mainly in the method used to fragment chromatin, says Chris Fry, who directs epigenetics and ChIP product development at the antibody company Cell Signaling Technology. Sonication-based methods subject the chromatin to harsh, denaturing conditions involving high heat, much detergent and strong shearing force that can damage antibody epitopes and chromatin integrity. These conditions tend to not agree with TFs and cofactors, he says. Labs can try lower detergent concentrations or cross-link with different formaldehyde concentrations. For successful ChIP, chromatin fragment length and integrity matter, says Fry. Many labs over-sonicate chromatin in the belief that shorter fragments are more suitable for ChIP-seq. “However, this over-sonication damages the chromatin and often results in complete loss of target protein and target loci enrichment,” he says.

Enzymatic digestion with micrococcal nuclease cuts the linker region between nucleosomes and fragments the chromatin more gently, and is likely more amenable to ChIP for TFs and cofactors, says Fry. One “pushback” he hears from customers is that the method is biased toward open chromatin. That is not the case, he says. Sonication might bias more toward open chromatin than enzymatic fragmentation. Labs have to choose between protein A and protein G beads for IP. Both work well for rabbit antibodies, but



A protein can be cross-linked to the DNA it interacts with and there can be non-specific cross-linking with nearby DNA sequences, says Velibor Savic.

Credit: J. Olday, Abcam

antibodies from mouse, goat and sheep bind better to protein G, he says. Labs should use beads tested and validated for ChIP-seq. “I always recommend using magnetic beads, not only for convenience, but also because they typically yield lower background than agarose or Sepharose beads,” he says.

Given differing methods for chromatin fragmentation and DNA library prep, says Fry, ChIP-seq experiments have an inherently high non-specific background signal with fragments that “come along for the ride.” They stick to the protein, beads or tubes even when the antibody is highly specific and binds strongly to the protein of interest. “This is why it is important to always include a ‘negative control’ in the experiment,” he says.

For ChIP-qPCR, he and his team recommend using normal IgG as a negative control to measure non-specific enrichment. For ChIP-seq, in his view, the preferred negative control is input chromatin. As a positive control for both ChIP-qPCR and ChIP-seq, labs can look at genomic regions known to bind the protein of interest or use the antibody against histone H3 as a ‘universal’ positive control, which will, he says, enrich for any genomic locus if the chromatin was intact and the IP worked as planned. “It always amazes me how many people perform their ChIP experiments without any positive or negative controls,” says Fry.

Antibody strategies

Enrichment in the wrong places, says Fry, is one reason that he recommends using a highly validated antibody for ChIP experiments. Labs can ask vendors about validation for ChIP or ChIP-seq and then closely study the data for levels of enrichment and the signal-to-noise in those assays. A “dirty” WB or incorrect localization in IF or immunohistochemistry (IHC) warns of potential non-specific binding in ChIP assays, he says. With histone-modification antibodies, Fry says labs should look for data that support specificity to the studied modification such as histone peptide array data. This assay comprehensively determines antibody

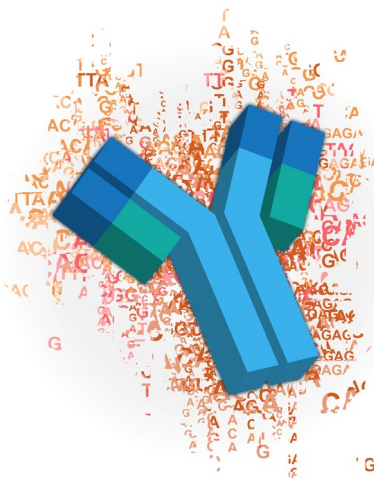
specificity to a modification site and helps reveal the effect of neighboring modifications on antibody binding. Using proprietary histone peptide arrays, Cell Signaling scientists test the specificity of the histone methyl-lysine, acetyl-lysine and methyl-arginine antibodies, and “we will provide data for any antibody upon customer request.” For planning a ChIP experiment, Fry recommends carefully selecting protocols and reagents. “Histone proteins, which are abundant and bind very strongly to DNA, are very forgiving and relatively easy to ChIP,” he says. TFs are less abundant, bind less stably to DNA and are more difficult to ChIP. Transcription cofactors that often do not directly contact DNA and may belong to large protein complexes are the most difficult target proteins to CHIP.

Commenting on DIP-seq flaws, Savic says epitopes tend to be larger targets than one DNA modification. Histone-modification antibodies may be specific to a particular modification, such as lysine tri-methylation, but the antibody recognizes around eight amino acids with the modified target among them. The same holds for nucleotide modifications. “It is safe to expect that the composition of the flanks would affect the binding of the antibody to a certain extent,” says Savic. Potentially, he says, using additional antibodies to the same target with different affinities to flanking sequences would alleviate the problem. Scientists shopping for ChIP antibodies, he says, should pick the antibody that will bind to the target in the cross-linked chromatin, recognize it in its native format and remain bound to beads and to the target throughout precipitation and wash. “Ideally, all four variables would be tested before embarking on a larger ChIP-seq experimental program,” he says.

More is not better

Validation should always include antibody titration. “Customers still tell me they need as much as 5 to 10 µg of antibody per immunoprecipitation in their ChIP assay,” says Fry. “This is simply not true. And more is not always better.” For many antibodies, both polyclonal and monoclonal, he and his team find the optimal range of antibody is 0.5–2 µg per IP. Adding too much antibody can decrease target enrichment.

Adding more antibody to troubleshoot will deliver variable results, says Savic. It might appear a promising way to raise signal; “however, it may also increase the signal rising from off-target interactions.” Off-target signals might blend with the background noise, but once the off-targets become more prominent and resemble



Avoiding false positives and correcting for off-target binding is best handled by normalizing DNA-modification enrichment to an IgG control, says the Nestor lab. Credit: A. Lentini, Linköping University

positive peaks, more antibody produces more false positives. Boosting a signal with a mix of antibodies raised against the same protein can help, he says. If raised against non-overlapping protein segments, these antibodies can interact with the target at the same time, which can translate to improved retention during IP-bead washing and better signal-to-noise ratios. But he advises caution: a polyclonal antibody is “usually raised only against the peptide of limited size, and therefore may in fact act equivalently to monoclonal in ChIP assays.” That’s when it helps to know the immunizing peptide to confirm coverage of the protein target’s distinctive part. The use of multiple antibodies does not apply to histone modifications, where only one antibody can bind per target because only one target exists per protein.

Validated or not

Some antibody companies such as Abcam and Cell Signaling Technology validate antibodies for ChIP. When buying a non-validated antibody, a lab needs to test and optimize the antibody for the ChIP assay, says Fry, for example, to titrate the antibody and look at enrichment of known target genes using ChIP-qPCR, to use controls and to consider testing more than one antibody to a target protein or protein complex. Just because an antibody is not validated for ChIP doesn’t mean it will fail ChIP, he says. “It may just not have been tested, and you will be the first to test it,” he says. A scientist can consult IP data for indications that an antibody is more likely to work in ChIP. If validated

for IF or IHC, an antibody might work in ChIP, since these assays involve cell fixation and epitope recognition of a protein in its more native conformation. If an antibody shows incorrect localization in IF or IHC or both, then cross-reactivity in a ChIP assay becomes more likely, he says. “Unfortunately, not all ‘validation’ is created equally,” he says. For example, some companies use the “ChIP Grade” label but don’t provide any validation data to support it.

Antibody next

Hiroshi Sasaki, a postdoctoral fellow at Harvard University’s Wyss Institute, battles noise and false positives in his single-cell super-resolution imaging of chromosomes. ChIP-seq is powerful for investigating chromatin state genome-wide and provides much data to statistically analyze, he says. Many labs have begun using ‘pulldown’ assays in single-cell biology, where improving signal-to-noise is tough. Most single-cell data suffer from non-specific signals that can’t be cancelled out as they can in population assays. Sasaki and his colleagues see non-specific signals in their single-molecule localization imaging with antibody labeling and mislocalized antibodies, but there is noise with oligo probes, too. All issues, including the specificity of antibodies and oligo probes, lot-to-lot stability, protocol-to-protocol optimization and experimental validation, “are getting more and more critical in modern biology,” he says.

Neither antibody selection nor validation is easy, says Nestor. There is also a limit to the number of validation experiments a researcher can be expected to do before the scientific community will accept results. DIP-seq caveats he and his group found were “not with the antibodies themselves, per se, but in how they controlled the experiments in which they used those antibodies,” says Nestor. Labs can consider antibody-free techniques and check independent replication in other labs. “But as we found, if the error is pervasive enough, it can hide in plain sight, through the consistency of the error across studies,” he says. “So I suppose, *caveat emptor*.”

Vivien Marx

Technology editor for Nature Methods.
e-mail: v.marx@us.nature.com

Published online: 28 March 2019
<https://doi.org/10.1038/s41592-019-0365-3>

References

1. Park, P. J. *Nat. Rev. Genet.* **10**, 669–680 (2009).
2. Meyer, C. A. & Liu, X. S. *Nat. Rev. Genet.* **11**, 709–721 (2014).
3. Jain, D. et al. *Nucleic Acids Res.* **43**, 6959–6968 (2015).
4. Lentini, A. et al. *Nat. Methods* **15**, 499–504 (2018).
5. Uhlen, M. et al. *Nat. Methods* **13**, 823–827 (2016).