JBC EDITORS' PICK HIGHLIGHT

A new approach for quantifying epigenetic landscapes

DOI 10.1074/jbc.H120.016430

Wolfgang Fischle*

From the Division of Biological and Environmental Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Edited by John M. Denu

ChIP-Seq is a widespread experimental method for determining the global enrichment of chromatin modifications and genome-associated factors. Whereas it is straightforward to compare the relative genomic distribution of these epigenetic features, researchers have also made efforts to compare their signal strength using external references for normalization. New work now suggests that these "spike-ins" could lead to inaccurate conclusions due to intrinsic issues of the methodology and instead calls for new criteria of experimental reporting that may permit internal standardization when certain parameters are fulfilled.

ChIP (1) has become a central technique for epigenetics research, revealing the occupation of defined regions of the genome with DNA-associated components (*e.g.* histones, histone modifications, transcription factors, chromatin modifiers, etc.). The experimental workflow begins with cross-linking of DNA-protein complexes for *in situ* fixation of a defined cellular state. Bulk chromatin is then isolated and fragmented, followed by enrichment of protein targets via immunoaffinity purification using specific antibodies. The corresponding enriched DNA fragments are recovered after that and measured. In classical ChIP, this is done by RT-PCR, targeting isolated regions of the genome. With the introduction of next-generation sequencing (NGS) to the methodology in 2007 (ChIP-Seq) (2), deducing the global genomic distribution of chromatin components became possible. The histograms (also known as epigenetic landscapes) determined by ChIP-Seq reflect the frequency of sequencing DNA fragments that were recovered by ChIP aligned with the cellular genome (Fig. 1*A*).

ChIP-Seq is technically relatively simple (*i.e.* it can be adopted by many laboratories), highly sensitive (due to intermediate PCR amplification and the power of NGS), and information-rich (due to standardized methods for processing large data sets). The approach has been instrumental for mapping functional genomic elements within several large scale international collaborative efforts (*e.g.* ENCODE (www.encodeproject.org) and IHEC (www.ihec-epigenomes.org)). However, it is not a perfect technique; problems can arise from imperfect reproducibility of cross-linking and chromatin fragmentation, during preparation of sequencing libraries and as sequencing artifacts, and—most of all—from incomplete or poorly defined antibody specificity.

ChIP-Seq histograms allow the comparison of the genomic distribution of different targets of antibodies (*i.e.* along the x coordinate), but what about the absolute enrichment (*i.e.* along the y axis)? If the signals of ChIP-Seq experiments were quantitative, the epigenetic states of different samples could be globally compared. As many parameters (*e.g.* variation of cell state, cell number, concentrations, cross-linking, intermediate amplification, sequencing conditions, etc.) impact the results of ChIP-Seq, there is no apparent direct correlation between the measured enrichment and the biological activity of a target in different experimental conditions. A seemingly straightforward solution to resolve this issue is normalization against known controls (3). It has been suggested to use internal referencing by executing two ChIP reactions simultaneously (against a target and a control) (4). Alternatively, external normalization by spiking ChIP-Seq experiments with foreign cells (5) or foreign chromatin (with single (6) or dual (7) antibody ChIP) has been proposed. Whereas these approaches should allow relative quantification as long as the spike-ins are kept constant between experiments, the use of recombinant nucleosomes as added references (8) could enable absolute quantification. All normalization procedures require that recoveries of the controls not be influenced by sample composition or experimental condition—an assumption that has not been directly examined. On the basis of conceptual considerations, Dickson *et al.* (9) now propose that quantitative comparison of ChIP-Seq experiments without spike-ins is possible, as long as certain experimental conditions are met. They call for a new standard in reporting ChIP-Seq experiments and show that spike-in quantification might produce misleading results in the context of antibodies with limited specificity and samples of high complexity (Fig. 1*B*).

The authors of the new study look at ChIP-Seq experiments as a closed reaction system (*i.e.* keeping reaction volumes as well as chromatin and antibody load constant). Using a series of physical and mathematical deductions, they infer that, under these experimental conditions, a proportionality factor α can be calculated that is derived from internal parameters of the biochemical, PCR amplification, and sequencing steps. These are the fractions of IP and input taken to library and NGS, the fraction of chromatin used as input, and the actual ratios of captured IP and input libraries. Because α establishes a real quantitative relation between the material used for the IP and the absolute number of defined NGS reads, Dickson *et al.* argue that these parameters should be reported for every ChIP-Seq experiment to facilitate data comparison.

To illustrate their approach, the authors first theoretically dissected and then experimentally analyzed a cell system where a histone modification epitope (H3K27me) is depleted by an inhibitor of the modifying enzyme. This is a highly representative scenario, as changes in the epigenetic landscape are frequently examined with regard to biological outcomes like cell

* For correspondence: Wolfgang Fischle, wolfgang.fischle@kaust.edu.sa.

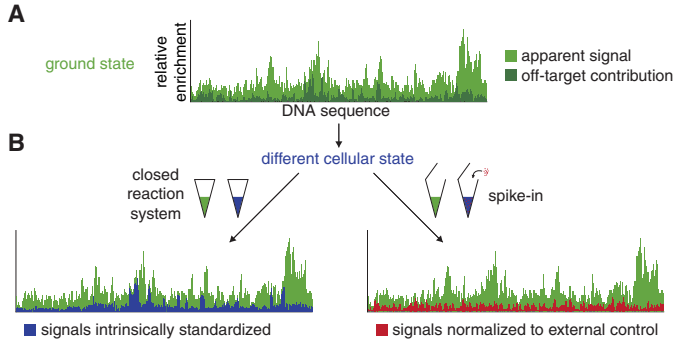

Figure 1. Comparative ChIP-Seq of different cellular states using different quantification schemes. *A*, a section of an epigenetic landscape as relative enrichment of DNA fragments (*y* axis) over the chromosomal location (*x* axis) is shown. The apparent signal generated by ChIP-Seq using a defined antibody might be a composite containing signals of target-specific and off-target recovery of DNA fragments. *B*, different procedures of quantification via intrinsic standardization (*left*, siQ-ChIP, reaction conditions need to be fully matched) and external normalization (*right*, spike-in, reaction conditions must ensure that the added material does not affect the experimental outcome) produce different epigenetic landscapes in cases where antibodies are cross-reactive.

differentiation and provide the basis for developing epigenetic drugs. Counterintuitively, Dickson *et al.* derive that, under certain ChIP-Seq conditions, enzyme inhibition could result in apparent increase in overall signal. This is because ChIP-Seq experiments are never carried out under perfect conditions (*i.e.* IP from a solution of a singular target or using antibodies with absolute selectivity). Instead, (i) cellular samples that serve as input for ChIP-Seq are highly complex (*i.e.* contain multiple epitopes that might be recognized and captured by the antibody with different affinities; also, general stickiness might occur); (ii) antibodies have limited specificity for their targets; and, even if all affinities were available for experimental corrections, (iii) the exact concentrations of on- and off-targets (reaction conditions) are not known.

The spike-in approach for comparing ChIP-Seq experiments introduces a normalization factor that should control for the capture efficiency of the antibody under different conditions. However, it does not provide any insights into the composition of the captured material, or "fractional composition." Dickson *et al.* illustrate the relation of this parameter to capture efficiency in an interactive online tool that allows users to simulate different theoretical ChIP-Seq conditions. According to these considerations, spike-ins allow accurate normalization only in a limited experimental window and can be deceiving in cases of antibody cross-reactivity and/or high abundance of off-targets. Conversely, with the introduction of the new proportionality factor α, it is possible to quantitatively compare ChIP-Seq reactions irrespective of the presence of off-targets and limited specificities of antibodies.

The new study provides an interesting conceptual framework for quantifying ChIP-Seq experiments, although the strict technical requirements (such as constant chromatin and antibody concentrations) define a narrow experimental window that might be challenging to fulfill. However, addressing the concerns raised by Dickson *et al.* regarding the limitations of normalization approaches in comparing epigenetic landscapes would also require more stringency to avoid artifacts and draw meaningful conclusions. The suggested reporting of additional experimental parameters should, in any case, help in evaluating different experiments, increasing reproducibility, and adding robustness to the method. It will be interesting to see whether combining spike-in approaches with the new considerations allows execution of comparative ChIP-Seq in broader experimental schemes. Also, it will need to be determined how other methods that do not rely on immunoaffinity capture but also make use of antibodies for mapping target genomic regions, such as CUT&RUN (10), can be conceptualized in a quantitative biology context and how such approaches compare with the current gold standard of ChIP-Seq. The work from Dickson *et al.* will, for sure, spark further discussions that will ultimately help the field improve the accuracy with which epigenetic landscapes among diverse biological conditions are defined.

Funding and additional information—Work in the laboratory of W. F. is supported by the King Abdullah University of Science and Technology (KAUST).

Conflict of interest—The author declares that he has no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: NGS, next-generation sequencing; ENCODE, Encyclopedia of DNA Elements; IHEC, International Human Epigenome Consortium; IP, immunoprecipitation.

References

1. Solomon, M. J., and Varshavsky, A. (1985) Formaldehyde-mediated DNA-protein crosslinking: a probe for *in vivo* chromatin structures. *Proc. Natl. Acad. Sci. U. S. A.* **82,** 6470–6474 CrossRef Medline
2. Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007) Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316,** 1497–1502 CrossRef Medline
3. Chen, K., Hu, Z., Xia, Z., Zhao, D., Li, W., and Tyler, J. K. (2015) The overlooked fact: fundamental need for spike-in control for virtually all genome-wide analyses. *Mol. Cell Biol.* **36,** 662–667 CrossRef Medline
4. Guertin, M. J., Cullen, A. E., Markowetz, F., and Holding, A. N. (2018) Parallel factor ChIP provides essential internal control for quantitative differential ChIP-seq. *Nucleic Acids Res.* **46,** e75 CrossRef Medline
5. Orlando, D. A., Chen, M. W., Brown, V. E., Solanki, S., Choi, Y. J., Olson, E. R., Fritz, C. C., Bradner, J. E., and Guenther, M. G. (2014) Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* **9,** 1163–1170 CrossRef Medline
6. Bonhoure, N., Bounova, G., Bernasconi, D., Praz, V., Lammers, F., Canella, D., Willis, I. M., Herr, W., Hernandez, N., Delorenzi, M., and CycliX Consortium (2014) Quantifying ChIP-seq data: a spiking method providing an internal reference for sample-to-sample normalization. *Genome Res.* **24,** 1157–1168 CrossRef Medline
7. Egan, B., Yuan, C. C., Craske, M. L., Labhart, P., Guler, G. D., Arnott, D., Maile, T. M., Busby, J., Henry, C., Kelly, T. K., Tindell, C. A., Jhunjhunwala, S., Zhao, F., Hatton, C., Bryant, B. M., *et al.* (2016) An alternative approach to ChIP-Seq normalization enables detection of genome-wide changes in histone H3 lysine 27 trimethylation upon EZH2 inhibition. *PLoS ONE* **11,** e0166438 CrossRef Medline
8. Grzybowski, A. T., Chen, Z., and Ruthenburg, A. J. (2015) Calibrating ChIP-Seq with nucleosomal internal standards to measure histone modification density genome wide. *Mol. Cell* **58,** 886–899 CrossRef Medline
9. Dickson, B. M., Tiedemann, R. L., Chomiak, A. A., Cornett, E. M., Vaughan, R. M., and Rothbart, S. B. (2020) A physical basis for quantitative ChIP-sequencing. *J. Biol. Chem.* **295,** 15826–15837 CrossRef Medline
10. Skene, P. J., and Henikoff, S. (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6,** e21856 CrossRef Medline

A new approach for quantifying epigenetic landscapes

Wolfgang Fischle

J. Biol. Chem. 2020, 295:15838-15839.
doi: 10.1074/jbc.H120.016430