

Video Article

Genome-wide Mapping of Protein-DNA Interactions with ChEC-seq in *Saccharomyces cerevisiae*

Sebastian Grünberg¹, Gabriel E. Zentner²¹Basic Sciences Division, Fred Hutchinson Cancer Research Center²Department of Biology, Indiana UniversityCorrespondence to: Gabriel E. Zentner at gzentner@indiana.eduURL: <https://www.jove.com/video/55836>DOI: [doi:10.3791/55836](https://doi.org/10.3791/55836)Keywords: Genetics, Issue 124, ChEC-seq, ChIP-seq, transcription, genome-wide, Mediator, *Saccharomyces cerevisiae*

Date Published: 6/3/2017

Citation: Grünberg, S., Zentner, G.E. Genome-wide Mapping of Protein-DNA Interactions with ChEC-seq in *Saccharomyces cerevisiae*. *J. Vis. Exp.* (124), e55836, doi:10.3791/55836 (2017).

Abstract

Genome-wide mapping of protein-DNA interactions is critical for understanding gene regulation, chromatin remodeling, and other chromatin-resident processes. Formaldehyde crosslinking followed by chromatin immunoprecipitation and high-throughput sequencing (X-ChIP-seq) has been used to gain many valuable insights into genome biology. However, X-ChIP-seq has notable limitations linked to crosslinking and sonication. Native ChIP avoids these drawbacks by omitting crosslinking, but often results in poor recovery of chromatin-bound proteins. In addition, all ChIP-based methods are subject to antibody quality considerations. Enzymatic methods for mapping protein-DNA interactions, which involve fusion of a protein of interest to a DNA-modifying enzyme, have also been used to map protein-DNA interactions. We recently combined one such method, chromatin endogenous cleavage (ChEC), with high-throughput sequencing as ChEC-seq. ChEC-seq relies on fusion of a chromatin-associated protein of interest to micrococcal nuclease (MNase) to generate targeted DNA cleavage in the presence of calcium in living cells. ChEC-seq is not based on immunoprecipitation and so circumvents potential concerns with crosslinking, sonication, chromatin solubilization, and antibody quality while providing high resolution mapping with minimal background signal. We envision that ChEC-seq will be a powerful counterpart to ChIP, providing an independent means by which to both validate ChIP-seq findings and discover new insights into genomic regulation.

Video Link

The video component of this article can be found at <https://www.jove.com/video/55836/>

Introduction

Mapping the binding sites of transcription factors (TFs), chromatin remodelers, and other chromatin-associated regulatory factors is key to understanding all chromatin-based processes. While chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq) approaches have been used to gain many important insights into genome biology, they have notable limitations. We recently introduced an alternative method, termed chromatin endogenous cleavage and high-throughput sequencing (ChEC-seq)¹, to circumvent these drawbacks.

ChIP-seq is most often performed with an initial formaldehyde crosslinking step (X-ChIP-seq) to preserve protein-DNA interactions. However, a number of recent studies have indicated that X-ChIP-seq captures transient or nonspecific protein-DNA interactions^{2,3,4,5,6,7,8}, giving rise to false positive binding sites. In addition, sonication, commonly used to fragment chromatin in X-ChIP-seq experiments, preferentially shears regions of open chromatin, leading to biased recovery of fragments from these regions^{9,10}. Sonication also yields a heterogeneous mixture of fragment lengths, ultimately limiting binding site resolution, though the addition of an exonuclease digestion step can greatly improve resolution^{11,12}. Native ChIP methods such as occupied regions of genomes from affinity-purified naturally isolated chromatin (ORGANIC)¹³ do not use crosslinking and fragment chromatin with micrococcal nuclease (MNase), alleviating potential biases associated with formaldehyde crosslinking and sonication. However, the solubility of many chromatin-bound proteins under the relatively mild conditions required for native chromatin extraction is poor, potentially leading to reduced dynamic range and/or false negatives¹⁴.

While various iterations of ChIP-seq are most commonly used for genome-wide mapping of protein-DNA interactions, several mapping techniques based on fusion of proteins of interest to various DNA-modifying enzymes have also been implemented. One such approach is DNA adenine methyltransferase identification (DamID)¹⁵, wherein a chromatin-binding protein of interest is genetically fused to Dam and this fusion is expressed in cells or animals, resulting in methylation of GATC sequences proximal to binding sites for the protein. DamID is advantageous in that it does not rely upon immunoprecipitation and so avoids crosslinking, antibodies, or chromatin solubilization. It is also performed *in vivo*. However, the resolution of DamID is limited to the kilobase scale and the methylating activity of the Dam fusion protein is constitutive. A second method based on enzymatic fusion is Calling Card-seq¹⁶, which employs fusion of a factor of interest to a transposase, directing site-specific integration of transposons. Like DamID, Calling Card-seq is not based on immunoprecipitation and thus has similar advantages, with the added benefit of increased resolution. However, Calling Card-seq may be limited by sequence biases of transposases and is also reliant on the presence of restriction sites close to transposon insertion sites.

A third enzymatic fusion method, developed in the Laemmli lab, is chromatin endogenous cleavage (ChEC)¹⁷. In ChEC, a fusion between a chromatin-associated protein and MNase is expressed in cells, and upon calcium addition to activate MNase, DNA is cleaved proximal to binding sites for the tagged factor (**Figure 1**). In conjunction with Southern blotting, ChEC has been used to characterize chromatin structure and protein binding at a number of individual loci in yeast^{17,18}, and has been combined with low-resolution microarray analysis to probe the interaction of nuclear pore components with the yeast genome¹⁹. ChEC offers benefits similar to DamID and Calling Card-seq, and its resolution is nearly single-base pair when analyzed by primer extension¹⁹. ChEC is also controllable: robust DNA cleavage by MNase depends on the addition of millimolar calcium, ensuring that MNase is inactive at the low free calcium concentrations observed in live yeast cells²⁰.

Previously, we postulated that combining ChEC with high-throughput sequencing (ChEC-seq) would provide high-resolution maps of TF binding sites. Indeed, ChEC-seq generated high-resolution maps of the budding yeast general regulatory factors (GRFs) Abf1, Rap1, and Reb1 across the genome¹. We have also successfully applied ChEC-seq to the modular Mediator complex, a conserved, essential global transcriptional coactivator²¹, expanding the applicability of ChEC-seq to megadalton-size complexes that do not directly contact DNA and may be difficult to map by ChIP-based methods. ChEC-seq is a powerful method both for independent validation of ChIP-seq results and generation of new insights into the regulation of chromatin-resident processes. Here, we present a step-by-step protocol for the implementation of this method in budding yeast.

Protocol

1. Generation of Yeast Strains

1. **Generate a yeast strain bearing the factor of interest tagged with MNase.**
 1. PCR amplify the MNase tagging cassette from the desired vector (**Table 1**) using the specified reaction mixture (**Table 2**) and cycling conditions (**Table 3**). Mix 5 μ L of the PCR reaction and 1 μ L of 6X DNA loading dye. Run each PCR aliquot on a 0.8% agarose gel at 120 V for 40 min. The expected product size is ~2.3 kb.
 2. Transform the amplified tagging cassette into the strain of choice using standard lithium acetate transformation²² and perform selection.
 3. Verify correct integration of the tagging cassette with colony PCR.
 1. Prepare the specified reaction mixture (**Table 4**) for the number of colonies to be screened. Keep on ice until needed.
 2. Pick a portion of each colony to be tested into the bottom of a PCR tube using a sterile toothpick or pipet tip.
 3. Microwave the picked colonies at high power for 1 min.
 4. Add 50 μ L of PCR mix (**Table 4**) to each microwaved colony and pipet up and down to mix. Perform PCR using the specified cycling conditions (**Table 5**).
 5. Mix the 50 μ L of PCR reaction and 10 μ L of 6X DNA loading dye directly in each PCR tube. Run 10 μ L of each sample on a 1% agarose gel at 120 V for 40 min. The expected product size is ~700 bp.

NOTE: If desired, verify appropriate expression of the MNase fusion protein via western blotting. A ~20.6 kilodalton shift in the molecular weight of the tagged protein is expected.
2. **Generate a free MNase strain by homology-based cloning of the promoter of the gene of interest into pGZ136 (Table 1) and transformation and selection as in step 1.1.2.**
 1. Alternatively, assemble the promoter of the gene of interest and 3xFLAG-MNase-SV40 NLS into a plasmid vector, transform and select as in step 1.1.2, and grow the strain in the appropriate selective condition for maintenance of the plasmid.

2. ChEC

1. In the afternoon/evening of the day prior to the ChEC experiment, inoculate 3 mL yeast-peptone-dextrose (YPD) or appropriate selective medium with a single colony of the strain bearing the MNase-tagged factor. Incubate this culture overnight (180 RPM shaking, 30 °C).
2. In the morning of the day of the ChEC experiment, dilute the overnight culture into 50 mL of medium to an optical density at 600 nm (OD₆₀₀) of 0.2 - 0.3. Incubate this culture (180 RPM shaking, 30 °C) until it reaches an OD₆₀₀ of 0.5-0.7. As the culture approaches the appropriate OD₆₀₀, set heat block or water bath to 30 °C and start thawing Buffer A additives (**Table 6**).
3. Prepare 5 mL of Buffer A plus additives (**Table 6**) per culture. Keep buffer A at RT during the procedure.
4. Prepare microfuge tubes containing 90 μ L of stop solution (**Table 7**) and 10 μ L of 10% SDS for each time point to be collected. For each new factor analyzed, take samples at 0 s, 30 s, 1 min, 2.5 min, 5 min, and 10 min.
5. Decant culture into a 50 mL conical tube and centrifuge at 1,500 x g and room temperature (RT) for 1 min.
6. Thoroughly resuspend cells in 1 mL of Buffer A and transfer the resuspended cells to a microfuge tube. Pellet resuspended cells in a microfuge at 1,500 x g and RT for 30 s.
7. Aspirate supernatant and thoroughly resuspend cells in 1 mL of Buffer A by pipetting up and down. Pellet resuspended cells in a microfuge at 1,500 x g and RT for 30 s. Repeat this step once.
8. Thoroughly resuspend cells in 570 μ L of Buffer A. Add 30 μ L of 2% digitonin (0.1% final concentration) to facilitate permeabilization of the cells, and invert to mix.
9. Permeabilize cells in heat block or water bath at 30 °C for 5 min.
10. Pipet the reaction up and down to ensure an even distribution of cells. Remove a 100 μ L aliquot of permeabilized cells as a negative control to a microfuge tube containing stop solution and SDS (step 2.4) and vortex briefly to mix.
11. Add 1.1 μ L of 1 M CaCl₂ (2 mM final concentration) to permeabilized cells to activate MNase and invert several times quickly or vortex briefly to mix. Immediately return the mixed reaction at 30 °C and start a timer.
12. At each time point to be collected, pipet the reaction up and down to ensure an even distribution of cells, then remove a 100 μ L aliquot of permeabilized cells to a microfuge tube containing stop solution and SDS and vortex briefly to mix. For each new factor analyzed, take 0 s (no CaCl₂ added), 30 s, 1 min, 2.5 min, 5 min, and 10 min time points.

NOTE: ChEC experiments with factors that rapidly cleave DNA at 30°C may be performed at lower temperatures to slow MNase cleavage kinetics.

13. Once all time points have been collected, add 4 µL of 20 mg/mL proteinase K to each sample, vortex briefly to mix, and incubate at 55 °C for 30 min.
14. Add 200 µL of 25:24:1 phenol:chloroform:isoamyl alcohol to each sample, vortex vigorously to mix, and centrifuge at maximum speed and RT in a microfuge for 5 min.
15. Remove each aqueous phase (~ 150 µL) to a new tube. Add 30 µg of glycogen and 500 µL of 100% ethanol. Vortex vigorously to mix and precipitate on dry ice for 10 min or until solution is viscous.
16. Pellet precipitated DNA at maximum speed and 4 °C in a microfuge for 10 min.
17. Decant supernatants and wash pellets with 1 mL of RT 70% ethanol.
18. Decant ethanol, removing the remainder by inverting and gently tapping the tubes on a paper towel. Briefly spin samples using a benchtop centrifuge and use a pipet to remove residual ethanol, taking care not to disturb the pellet. Air-dry pellets for at RT for 5 min.
19. While pellets are drying, make a master mix to resuspend dried pellets in 29 µL of Tris, pH 8.0 + 1 µL of 10 mg/mL RNase A each. Digest RNA at 37 °C for 20 min.
20. Mix 1 µL of 6X DNA loading dye with 5 µL of each sample and run on a 1.5% agarose gel at 120 V for 40 min.

3. Size Selection

NOTE: The goal of size selection is to remove multi-kilobase fragments of genomic DNA from the sample to be sequenced and enrich fragments of ~150 bp (approximately the size of nucleosomal DNA) or smaller. In sequencing data, fragments <400 bp are enriched, with a notable peak around the size of nucleosomal DNA (~150 bp) and a peak or broad distribution of subnucleosomal fragments.

1. To each sample, add 75 µL of solid-phase reversible immobilization (SPRI) beads in a polyethylene glycol (PEG) solution²³ and pipet up and down 10 times to mix. Incubate the bead:DNA mixture at RT for 5 min.
2. During the SPRI bead incubation, prepare microfuge tubes containing 96 µL of 10 mM Tris, pH 8.0 and 4 µL of 5 M NaCl for each sample.
3. Place tubes in a magnetic rack to collect beads on the side of the tube. Allow the beads to collect on the side of the tube for 2 min.
4. Transfer each supernatant to a new tube containing Tris and NaCl (step 3.2).
5. Add 200 µL of 25:24:1 phenol:chloroform:isoamyl alcohol to each sample, vortex to mix, and centrifuge at maximum speed and RT in a microfuge for 5 min.
6. Remove each aqueous phase to a new tube. Add 30 µg of glycogen and 500 µL of 100% ethanol. Precipitate DNA on dry ice for 10 min or until solution is viscous.
7. Pellet precipitated DNA at maximum speed and 4 °C in a microfuge for 10 min.
8. Decant supernatants and wash pellets with 1 mL of 70% ethanol.
9. Decant ethanol, removing the remainder by inverting and gently tapping the tubes on a paper towel. Briefly spin samples using a benchtop centrifuge and use a pipet to remove residual ethanol, taking care not to disturb the pellet. Air-dry pellets for at RT for 5 min.
10. Resuspend dried pellets in 25 µL of Tris, pH 8.0. Determine sample concentration using a high-sensitivity assay. For TF ChEC experiments, 10-100 ng DNA is routinely recovered after size selection.
11. Prepare sequencing libraries. Library preparation protocols that do not rely upon size selection, as previously described²⁴, are desirable to allow sequencing of a large range of fragment sizes (25-500 bp).
12. Sequence libraries in paired-end mode. A minimum of 25 bases of sequence on each end is needed for high quality read mapping. Between 1 and 2 million reads provides sufficient coverage for a ChEC sample.

NOTE: Analysis of ChEC-seq data is a complex procedure and beyond the scope of this article. Detailed information on data analysis can be found in previous publications^{1,21}, and the custom scripts used for analysis are available on GitHub (<https://github.com/zentnerlab/chec-seq>). HOMER²⁵ (<http://homer.salk.edu>) can also be used for command-line analysis of ChEC-seq data.

Representative Results

In the case of a successful ChEC experiment, analysis of DNA by agarose gel electrophoresis will reveal a calcium-dependent increase in DNA fragmentation over-time, as indicated by smearing and eventual complete digestion of genomic DNA. In some cases, a ladder of bands similar to that seen with a traditional MNase digestion is observed after extended digestion. This is the case for ChEC analysis of Reb1, a general regulatory factor that binds nucleosome-depleted regions (NDRs) (**Figure 2**). We have found, in the case of GRFs, extended digestion leads to loss of signal at rapidly cleaved binding sites¹. Ideally, a sample with a reduction in the size of the genomic DNA band that displays high-molecular weight smearing, such as the 30 s and 1 min time points for Reb1 ChEC, will be used for sequencing to avoid time-dependent loss of signal.

Visualization of fragment ends from a Reb1 ChEC-seq experiment reveals robust enrichment of Reb1 over background derived from both the Reb1-MNase strain without calcium added and the free MNase strain after an equal duration of free calcium treatment (30 s) (**Figure 3**). Similarly, we observe substantial DNA cleavage by a fusion of the Mediator subunit Med8 with MNase, but not free MNase driven by the *MED8* promoter, 1 min after calcium addition (**Figure 3**).

To compare Reb1 ChEC-seq data to ChIP-seq, we obtained a list of 1,991 Reb1 peaks determined by ORGANIC¹³. These peaks were motif-centered with the average fragment end count at each base position in a 100 bp window around the motif midpoint. We observed a striking asymmetry in cleavage, with the majority of fragment ends mapping to the upstream side of the motif (**Figure 4**).

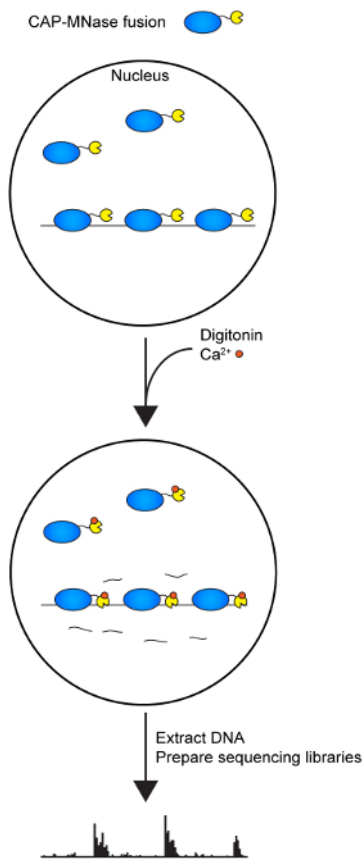


Figure 1: Schematic of the ChEC-seq Method. A chromatin-associated protein (CAP)-MNase fusion is expressed in yeast cells. The protein binds to DNA but does not generate cleavage above background levels due to the very low free calcium in the nucleus. Upon permeabilization of cells with digitonin and addition of millimolar calcium, CAP-MNase fusions bound to the genome cleave DNA, releasing small fragments. These fragments are then purified, sequenced, and mapped back to the genome, giving peaks of fragment ends proximal to binding sites for the CAP-MNase fusion. [Please click here to view a larger version of this figure.](#)

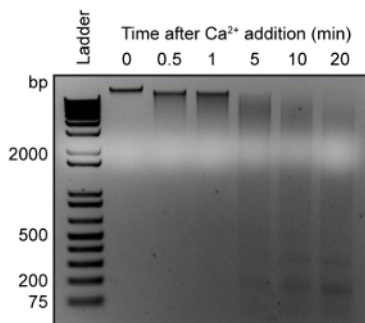


Figure 2: Agarose Gel Electrophoresis of DNA from a Reb1 ChEC Experiment. A 5 μ L aliquot of DNA from each ChEC time point was analyzed on a 1.5% TAE-agarose gel prior to size selection. This shows progressive digestion of genomic DNA by the Reb1-MNase fusion. [Please click here to view a larger version of this figure.](#)

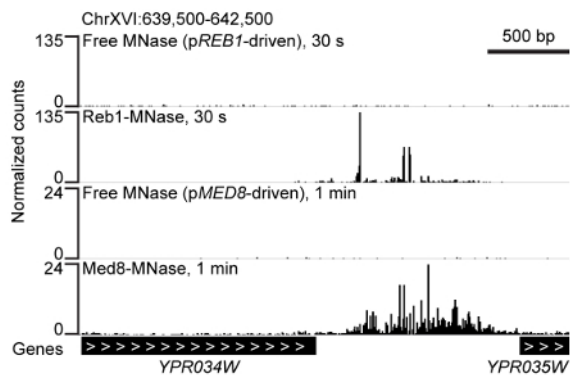


Figure 3: Genome Browser Snapshots of Reb1-MNase and Free MNase ChEC-seq Experiments. IGV views of fragment end signal for Reb1 and free MNase ChEC-seq 30 s after calcium addition and Med8 and free MNase ChEC-seq 1 min after calcium addition along a representative segment of the yeast genome. Datasets were normalized by dividing the number of fragment ends mapped to each base position by the total number of fragment ends mapped and multiplying by the total number of bases mapped. [Please click here to view a larger version of this figure.](#)

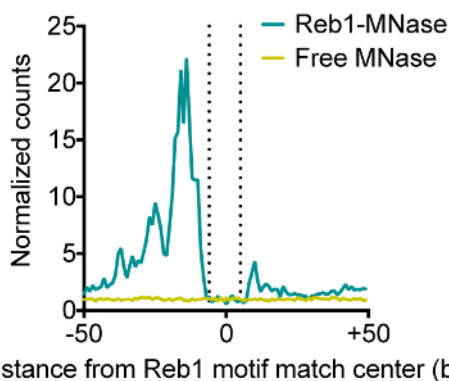


Figure 4: Enrichment of Reb1-MNase-released Fragment Ends around Reb1 ORGANIC Sites. Average plot of 30 s Reb1 and free MNase ChEC-seq fragment end signal around 1,991 Reb1 motifs determined by ORGANIC¹³. Data were normalized as in **Figure 3**. [Please click here to view a larger version of this figure.](#)

Plasmid	Yeast selectable marker	Notes	Addgene plasmid number
pGZ108	kanMX6	3xFLAG-MNase tagging, 33 aa linker	70231
pGZ109	HIS3MX6	3xFLAG-MNase tagging, 33 aa linker	70232
pGZ110	TRP1	3xFLAG-MNase tagging, 33 aa linker	70233
pGZ136	URA3	Expresses 3xFLAG-MNase-SV40 NLS under the control of the <i>REB1</i> promoter	72273
pGZ173	kanMX6	MNase tagging, 8 aa linker	70234

Table 1: Details of ChEC Plasmids. All tagging vectors are based on pFA6a vectors and so are compatible with the commonly used F2/R1 tagging primer pairs. The tagging cassette consists of a linker of the indicated length, a 3xFLAG epitope to facilitate detection by western blotting (except in the case of pGZ173, where the linker has been shortened to remove the 3xFLAG tag), the mature chain of MNase (GenBank P00644, aa 83 - 231), and the indicated selectable marker.

Reagent	Volume	[Final]
5x PCR buffer	10 mL	1x (2 mM MgCl ₂)
10 mM dNTP mix (2.5 mM each dNTP)	1 mL	200 mM (50 mM each dNTP)
10 mM F2 primer	2.5 mL	0.5 mM
10 mM R1 primer	2.5 mL	0.5 mM
pGZ108/109/110/172 (1-5 ng/mL)	1 mL	
2 U/μL hot start high-fidelity polymerase	0.5 mL	1 U
ddH ₂ O	32.5 mL	

Table 2: Reaction Mixture for PCR Amplification of 3xFLAG-MNase Tagging Cassettes. F2/R1 primer sequences can be found at <http://yeastgfp.yeastgenome.org/yeastGFPOligoSequence.txt>.

Temperature (°C)	Time	Cycles
98	30 s	1
98	10 s	25
55	30 s	25
72	1 min 15 s	25
72	2 min	1
10	Forever	Hold

Table 3: Thermal Cycling Conditions for PCR Amplification of 3xFLAG-MNase Tagging Cassettes. Incubation temperature and extension time may need to be adjusted based on the DNA polymerase used.

Reagent	Volume	[Final]
10x Taq buffer	5 mL	1x (1.5 mM MgCl ₂)
10 mM dNTP mix (2.5 mM each dNTP)	1 mL	200 mM (50 mM each dNTP)
10 mM Check primer	1 mL	0.2 mM
10 mM MNase-R or negative control primer	1 mL	0.2 mM
5 U/mL Taq polymerase	0.5 mL	2.5 U
ddH ₂ O	41.5 mL	

Table 4: Reaction Mixture for Colony PCR Confirmation of MNase Tagging. Check primer sequences can be found at <http://yeastgfp.yeastgenome.org/yeastGFPOligoSequence.txt>. The check primer is used as the forward primer along with a reverse primer within MNase (5'-TTGTGCAGCTTCTTGGTAC-3') and a reverse primer #500 base pairs (bp) downstream of the respective gene as a negative control.

Temperature (°C)	Time	Cycles
95	5 min	1
95	30 s	35
55	30 s	35
68	1 min	35
68	5 min	1
10	Forever	Hold

Table 5: Thermal Cycling Conditions for Colony PCR Confirmation of MNase Tagging. Incubation temperature and extension time may need to be adjusted based on the DNA polymerase used.

Reagent	Volume	[Final]
1 M Tris, pH 7.5	1.5 mL	15 mM
1 M KCl	8 mL	80 mM
0.2 M EGTA	50 mL	0.1 mM
ddH ₂ O	Fill to 100 mL	

Table 6: Recipe for Buffer A. Prior to use, add half of a protease inhibitor tablet, 50 μL of 100 mM PMSF (1 mM final), 5 μL of 200 mM spermine (0.2 mM final), and 2.5 μL of 1 M spermidine (0.5 mM final) per 5 mL of solution.

Reagent	Volume	[Final]
5 M NaCl	8 mL	400 mM
0.5 M EDTA	4 mL	20 mM
0.2 M EGTA	2 mL	4 mM
ddH ₂ O	Fill to 100 mL	

Table 7: Recipe for 2x Stop Solution. Combine 90 μ L of stop solution with 10 μ L of 10% SDS in a microfuge tube for each time point to be taken (see step 2.4).

Discussion

We have shown that ChEC can map diverse classes of yeast proteins on chromatin, and anticipate that it will be broadly applicable to different families of TFs and other chromatin-binding factors in yeast. ChEC-seq is advantageous in that it does not require crosslinking, chromatin solubilization, or antibodies. Thus, ChEC avoids artifacts potentially present in X-ChIP-seq, such as the hyper-ChIPable artifact^{3,4}, and native ChIP, such as false negatives due to incomplete protein solubilization¹⁴. The major drawback of ChEC-seq, as with all enzymatic fusion methods, is the requirement for a fusion protein. This can be achieved quickly in budding yeast, but is more laborious in metazoans. Another limitation of ChEC-seq is that it cannot be implemented as-is for profiling of histone modifications. However, a modification-binding domain could be fused to MNase and expressed, enabling histone modification mapping with ChEC-seq. In this vein, ChIP-seq using epitope-tagged histone modification-binding domains has been used to map genome-wide patterns of histone modification²⁶.

The use of a free MNase control is important in establishing the specificity of ChEC-seq results. The free MNase strain bears MNase tagged with 3xFLAG and a simian virus 40 nuclear localization signal (SV40 NLS) under the control of the endogenous promoter for the gene of interest. It is conceptually similar to the unfused Dam control used in DamID studies²⁷ and controls for nonspecific cleavage due to chromatin accessibility. As a free MNase control for TF ChEC-seq, we integrated 3xFLAG-MNase-SV40 NLS driven by the *REB1* promoter at the *ura3* locus¹. For Mediator ChEC-seq, we expressed 3xFLAG-MNase-SV40 NLS under the control of the *MED8* promoter in a non-integrating plasmid vector maintained in selective medium²¹. In both cases, very little cleavage by free MNase was observed. Thus, a single control strain wherein 3xFLAG-MNase-SV40 NLS is driven by a relatively robust promoter should be appropriate for most experiments.

When planning a ChEC experiment, structural consideration of the protein of interest may be important. For instance, we found that appending MNase to the C-terminus of Reb1 gave an asymmetric cleavage pattern around Reb1 motifs, while expression of Reb1 with N-terminal MNase gave a symmetric cleavage pattern¹. We attribute these different cleavage patterns to the structure of Reb1. The DNA binding domain of Reb1 is located at its C-terminus; thus, the C-terminus is structured and close to DNA, limiting the range of motion of MNase and only allowing C-terminal cleavage. In contrast, the C-terminus of Abf1 is relatively unstructured and may thus act to increase to reach of MNase, allowing cleavage on both sides of Abf1 binding sites. In the case of Rap1, another yeast GRF, we found that shortening the linker between the C-terminus and MNase from 33 to 8 aa was sufficient to severely reduce cleavage, perhaps due to the proposed distance of the C-terminal portion of Rap1 from DNA²⁸. Such structural consideration is also likely to be important when attempting to map the binding of proteins present in large complexes. For our ChEC-seq analysis of Mediator binding to the yeast genome²¹, we selected subunits whose C-termini were structurally predicted to be exposed, rather than buried within the complex. Of the three subunits fused to MNase, one did not robustly cleave DNA, suggesting that either steric constraints, or interactions with other DNA binding factors attenuated DNA cleavage.

We found that ChEC-seq detects 4-12 times more peaks for the yeast GRFs Abf1, Rap1, and Reb1 than various ChIP approaches when all time points were considered together¹. These sites could be divided into two classes based on MNase cleavage kinetics. Analysis of ChEC-seq data for yeast GRFs revealed two temporally distinct classes of binding sites¹. The first, termed 'fast,' displayed maximal cleavage <1 min after calcium addition and generally containing robust matches to known consensus motifs. The second, termed 'slow,' took several minutes to reach appreciable levels of cleavage and were depleted of motif matches. Both classes of sites were, on average, enriched for binding in ChIP datasets, though they were not necessarily called as peaks in those ChIP studies. The majority of sites for a given factor were slow sites. We speculate that fast sites are representative of high-affinity transcription factor binding sites, while slow sites represent loci transiently sampled during binding site scanning. The observation of two classes of binding sites separated by the kinetics of MNase cleavage may explain the observation in many ChIP-seq datasets that a large number of peaks for a given factor with a well-established DNA binding specificity do not contain a consensus motif²⁹: formaldehyde crosslinking, performed for 10-15 min in most ChIP protocols, can repeatedly capture transient or opportunistic interactions of a factor with the genome, leading to inflation of signal. Indeed, comparison of ChIP-exo and live cell imaging suggests that this is the case². Notably, we did not observe such a dramatic time-dependent decrease in Mediator ChEC-seq signals²¹, with Mediator ChEC-seq signal relatively constant from 30 s to 20 min. Given these observations, we recommend performing short-duration ChEC-seq to recover high-confidence binding sites.

We anticipate that ChEC-seq can be adapted to non-yeast systems. The concentration of free calcium in unstimulated mammalian cells is 50-300 nM^{30,32}, far below the threshold for efficient MNase activation. The limiting step for establishment of ChEC-seq in metazoan systems is thus the generation of fusion proteins, which remains far more laborious in metazoan than yeast cells. However, endogenous tagging of metazoan genes has been greatly facilitated by the advent of CRISPR-based genome engineering³³, and so this limitation should be readily surmounted. Alternatively, MNase fusion proteins could be expressed at low levels from plasmids. This approach has been quite successful for DamID in *Drosophila*^{34,35} and human³⁶ cells and thus might also facilitate ChEC-seq in metazoan cells.

Disclosures

The authors have nothing to disclose.

Acknowledgements

We thank Moustafa Saleh and Jay Tourigny for critical reading of the manuscript and Steven Hahn and Steven Henikoff for mentorship and support during the development of ChEC-seq and its application to the Mediator complex. S.G. is supported by NIH grants R01GM053451 and R01GM075114 and G.E.Z. is supported by Indiana University startup funds.

References

- Zentner, G. E., Kasinathan, S., Xin, B., Rohs, R., & Henikoff, S. ChEC-seq kinetics discriminates transcription factor binding sites by DNA sequence and shape in vivo. *Nat Commun.* **6** 8733 (2015).
- Chen, J. *et al.* Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. *Cell.* **156** (6), 1274-1285 (2014).
- Teytelman, L., Thurtle, D. M., Rine, J., & van Oudenaarden, A. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proc Natl Acad Sci USA.* **110** (46), 18602-18607 (2013).
- Park, D., Lee, Y., Bhupindersingh, G., & Iyer, V. R. Widespread Misinterpretable ChIP-seq Bias in Yeast. *PLoS ONE.* **8** (12), e83506 (2013).
- Jain, D., Baldi, S., Zabel, A., Straub, T., & Becker, P. B. Active promoters give rise to false positive 'Phantom Peaks' in ChIP-seq experiments. *Nucleic Acids Res.* (2015).
- Krebs, W. *et al.* Optimization of transcription factor binding map accuracy utilizing knockout-mouse models. *Nucleic Acids Res.* **42** (21), 13051-13060 (2014).
- Paul, E., Zhu, Z. I., Landsman, D., & Morse, R. H. Genome-Wide Association of Mediator and RNA Polymerase II in Wild-Type and Mediator Mutant Yeast. *Mol Cell Biol.* **35** (1), 331-342 (2015).
- Worsley Hunt, R., & Wasserman, W. W. Non-targeted transcription factors motifs are a systemic component of ChIP-seq datasets. *Genome Biol.* **15** (7), 412 (2014).
- Vega, V. B., Cheung, E., Palanisamy, N., & Sung, W.-K. Inherent Signals in Sequencing-Based Chromatin-ImmunoPrecipitation Control Libraries. *PLoS ONE.* **4** (4), e5241 (2009).
- Teytelman, L. *et al.* Impact of Chromatin Structures on DNA Processing for Genomic Analyses. *PLoS ONE.* **4** (8), e6700 (2009).
- Rhee, H. S., & Pugh, B. F. Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution. *Cell.* **147** (6), 1408-1419 (2011).
- He, Q., Johnston, J., & Zeitlinger, J. ChIP-nexus enables improved detection of in vivo transcription factor binding footprints. *Nat Biotech.* **33** (4), 395-401 (2015).
- Kasinathan, S., Orsi, G. A., Zentner, G. E., Ahmad, K., & Henikoff, S. High-resolution mapping of transcription factor binding sites on native chromatin. *Nat Methods.* **11** (2), 203-209 (2014).
- Skene, P. J., Hernandez, A. E., Groudine, M., & Henikoff, S. The nucleosomal barrier to promoter escape by RNA polymerase II is overcome by the chromatin remodeler Chd1. *eLife.* **3** e02042 (2014).
- Aughey, G. N., & Southall, T. D. Dam it's good! DamID profiling of protein-DNA interactions. *Wiley Interdiscip Rev Dev Biol.* **5** (1), 25-37 (2016).
- Wang, H., Mayhew, D., Chen, X., Johnston, M., & Mitra, R. D. Calling Cards enable multiplexed identification of the genomic targets of DNA-binding proteins. *Genome Res.* **21** (5), 748-755 (2011).
- Schmid, M., Durussel, T., & Laemmli, U. K. ChIC and ChEC: Genomic Mapping of Chromatin Proteins. *Mol Cell.* **16** (1), 147-157 (2004).
- Merz, K. *et al.* Actively transcribed rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules. *Genes Dev.* **22** (9), 1190-1204 (2008).
- Schmid, M. *et al.* Nup-PI: The Nucleopore-Promoter Interaction of Genes in Yeast. *Mol Cell.* **21** (3), 379-391 (2006).
- Dunn, T., Gable, K., & Beeler, T. Regulation of cellular Ca²⁺ by yeast vacuoles. *J Biol Chem.* **269** (10), 7273-7278 (1994).
- Grünberg, S., Henikoff, S., Hahn, S., & Zentner, G. E. Mediator binding to UASs is broadly uncoupled from transcription and cooperative with TFIID recruitment to promoters. *The EMBO Journal.* **35** (22), 2435-2446 (2016).
- Gietz, R. D., & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protocols.* **2** (1), 31-34 (2007).
- Rohland, N., & Reich, D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22** (5), 939-946 (2012).
- Orsi, G. A., Kasinathan, S., Zentner, G. E., Henikoff, S., & Ahmad, K. Mapping regulatory factors by immunoprecipitation from native chromatin. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.].* **110** 21-25 (2015).
- Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell.* **38** (4), 576-589 (2010).
- Kungulovski, G. *et al.* Application of histone modification-specific interaction domains as an alternative to antibodies. *Genome Res.* **24** (11), 1842-1853 (2014).
- Vogel, M. J., Peric-Hupkes, D., & van Steensel, B. Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nat. Protocols.* **2** (6), 1467-1478 (2007).
- Feeser, E. A., & Wolberger, C. Structural and Functional Studies of the Rap1 C-Terminus Reveal Novel Separation-of-Function Mutants. *J Mol Biol.* **380** (3), 520-531 (2008).
- Neph, S. *et al.* An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature.* **489** (7414), 83-90 - supplementary-information (2012).
- Mountjoy, K. G., Kong, P. L., Taylor, J. A., Willard, D. H., & Wilkison, W. O. Melanocortin receptor-mediated mobilization of intracellular free calcium in HEK293 cells. *Physiol Genomics.* **5** (1), 11-19 (2001).
- Seurin, D., Lombet, A., Babajko, S., Godeau, F., & Ricort, J.-M. Insulin-Like Growth Factor Binding Proteins Increase Intracellular Calcium Levels in Two Different Cell Lines. *PLoS ONE.* **8** (3), e59323 (2013).
- Grienberger, C., & Konnerth, A. Imaging Calcium in Neurons. *Neuron.* **73** (5), 862-885 (2012).
- Sander, J. D., & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotech.* **32** (4), 347-355. (2014).

34. Filion, G. J. *et al.* Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in *Drosophila* Cells. *Cell*. **143** (2), 212-224 (2010).
35. van Bommel, J.G. *et al.* A Network Model of the Molecular Organization of Chromatin in *Drosophila*. *Mol Cell*. **49** (4), 759-771 (2013).
36. Izzo, A. *et al.* The Genomic Landscape of the Somatic Linker Histone Subtypes H1.1 to H1.5 in Human Cells. *Cell Rep*. **3** (6), 2142-2154 (2013).