Video Article

Adaptation of Hybridization Capture of Chromatin-associated Proteins for Proteomics to Mammalian Cells

Hector Guillen-Ahlers[1,2], Prahlad K. Rao[1], Danu S. Perumalla[1], Maria J. Montoya[1], Avinash Y.L. Jadhav[1], Michael R. Shortreed[3], Lloyd M. Smith[3], Michael Olivier[1,2]

[1]Department of Genetics, Texas Biomedical Research Institute

[2]Department of Internal Medicine-Molecular Medicine, Wake Forest University School of Medicine

[3]Department of Chemistry, University of Wisconsin

Correspondence to: Michael Olivier at molivier@wakehealth.edu

URL: https://www.jove.com/video/57140
DOI: doi:10.3791/57140

Keywords: Genetics, Issue 136, HyCCAPP, proteomics, DNA-protein interactions, chromatin, mass spectrometry, NFκB, DUSP3, DUSP5

Date Published: 6/1/2018

Citation: Guillen-Ahlers, H., Rao, P.K., Perumalla, D.S., Montoya, M.J., Jadhav, A.Y., Shortreed, M.R., Smith, L.M., Olivier, M. Adaptation of Hybridization Capture of Chromatin-associated Proteins for Proteomics to Mammalian Cells. *J. Vis. Exp.* (136), e57140, doi:10.3791/57140 (2018).

Abstract

The hybridization capture of chromatin-associated proteins for proteomics (HyCCAPP) technology was initially developed to uncover novel DNA-protein interactions in yeast. It allows analysis of a target region of interest without the need for prior knowledge about likely proteins bound to the target region. This, in theory, allows HyCCAPP to be used to analyze any genomic region of interest, and it provides sufficient flexibility to work in different cell systems. This method is not meant to study binding sites of known transcription factors, a task better suited for Chromatin Immunoprecipitation (ChIP) and ChIP-like methods. The strength of HyCCAPP lies in its ability to explore DNA regions for which there is limited or no knowledge about the proteins bound to it. It can also be a convenient method to avoid biases (present in ChIP-like methods) introduced by protein-based chromatin enrichment using antibodies. Potentially, HyCCAPP can be a powerful tool to uncover truly novel DNA-protein interactions. To date, the technology has been predominantly applied to yeast cells or to high copy repeat sequences in mammalian cells. In order to become the powerful tool we envision, HyCCAPP approaches need to be optimized to efficiently capture single-copy loci in mammalian cells. Here, we present our adaptation of the initial yeast HyCCAPP capture protocol to human cell lines, and show that single-copy chromatin regions can be efficiently isolated with this modified protocol.

Video Link

The video component of this article can be found at https://www.jove.com/video/57140/

Introduction

During the past decade, there has seen a dramatic improvement in sequencing technologies, allowing the study of a wide range of genomes in large numbers of samples, and with astonishing resolution. The Encyclopedia of DNA Elements (ENCODE) Consortium, a large-scale multi-institutional effort spearheaded by the National Human Genome Research Institute of the National Institutes of Health, has provided insights into how individual transcription factors and other regulatory proteins bind to and interact with the genome. The initial effort characterized specific DNA-protein interactions, as assessed by Chromatin immunoprecipitation (ChIP) for over 100 known DNA-binding proteins[1]. Alternative methods such as DNase footprinting[2] and formaldehyde assisted isolation of regulatory elements (FAIRE)[3] have also been used to locate specific regions of the genome interacting with proteins, but with the obvious limitation that these experimental approaches do not identify the interacting proteins. Despite the extensive efforts over the past years, no technology has emerged that efficiently allows the comprehensive characterization of protein-DNA interactions in chromatin, and the identification and quantification of chromatin-associated proteins.

To address this need, we developed a novel approach which we termed as Hybridization Capture of Chromatin-Associated Proteins for Proteomics (HyCCAPP). Initially developed in yeast[4,5,6], the approach isolates crosslinked chromatin regions of interest (with bound proteins) using sequence-specific hybridization capture. After isolation of the protein-DNA complexes, approaches such as mass spectrometry can be used to characterize the set of proteins bound to the sequence of interest. Thus, HyCCAPP can be considered as a non-biased approach to uncover novel DNA-protein interactions, in the sense that it does not rely on antibodies and it is completely agnostic about the proteins that might be found. There are other approaches capable to uncover novel DNA-interacting proteins[7], but most rely on ChIP-like methods[8,9,10], plasmid insertions[11,12,13,14], or regions with high copy numbers[15]. In contrast, HyCCAPP can be applied to multi- and single-copy regions, and it does not require any prior information about the proteins in the region. In addition, while some of the methods mentioned above have valuable features, notably avoiding the need for DNA-protein crosslinking reactions, the unique feature of HyCCAPP is that it can be applied to single-copy regions in unmodified cells, and without any prior knowledge about putative binding proteins, or available antibodies.

At this point, HyCCAPP has predominantly been applied to the analysis of various genomic regions in yeast[4,5,6], and was recently used to analyze protein-DNA interactions in alpha-satellite DNA, a repeat region in the human genome[16]. As part of our ongoing work, we have adapted the hybridization capture approach initially developed for yeast chromatin to be applicable to the analysis of human cells, and present here a

modified protocol that allows the selective capture of single-copy target regions in the human genome with efficiencies similar to our initial studies in yeast. This new optimized protocol now allows the adaptation and utilization of the technology to interrogate protein-DNA interactions across the human genome, using mass spectrometry or other analytical approaches.

It is important to emphasize that the HyCCAPP method is meant for the analysis of specific target regions and is not yet suitable for genome-wide analyses. The technology is especially useful when dealing with regions for which there is scarce information about interacting proteins, or when a more comprehensive in-depth analysis of interacting proteins at a specific genome locus is desired. HyCCAPP is meant to uncover DNA-binding proteins but not characterize accurately the specific protein binding sites in genomic DNA. In its current implementation, the methodology does not provide information about the DNA binding sequences or motifs for individual proteins. Therefore, it nicely complements existing technologies such as FAIRE, and may allow the identification of novel binding proteins in genomic regions identified by an initial FAIRE analysis.

Protocol

1. Capture Oligonucleotide Design

1. **Design a panel of oligonucleotides to be used during the hybridization capture of the target region/s.**
    1. Aim to design 4–8 oligonucleotides per target region, but as a minimum, design at least one oligonucleotide targeting each end of the target region.
    2. If the target sequence is long (>500 base pairs), design the oligonucleotides as spread out as possible to assure effective enrichment of chromatin across the entire target region.
    NOTE: We have observed optimal captures with 4–8 oligonucleotides. Even though this process can work well with more oligonucleotides, it is not uncommon to observe a decrease in enrichment yields when too many oligonucleotides are used.

2. Design oligonucleotides with similar melting temperatures, and make sure that they do not form strong hairpins nor have strong interactions (taking into account that hybridizations will be carried out at 42 ºC).
NOTE: If a toehold elution is chosen (4.12.3), an additional external 8 base sequence must be added to the oligonucleotide and will require complementary 'release' oligonucleotides at the time of the elution[17]. Several commercial software tools can be used to select and analyze the oligonucleotides. 25 mers have shown good results, but depending on the genome size and sequence, capture oligonucleotides can contain somewhere between 20–80 bases. Typically, capture oligonucleotides have a melting temperature close to 62 ºC, but that can change depending on their length. Consistent melting temperatures and/or oligonucleotide length throughout the capture mix can help avoid unwanted biases in enrichment.
3. Design the capture oligonucleotides with a biotin moiety (preferably separated by a spacer from the capture sequence) on the 3' end.

2. Cell Culture

1. **Grow human lymphoblastoid cells (such as GM12878 shown here) or other cell lines in RPMI 1640 media supplemented with 1% penicillin-streptomycin, 2 mM L-glutamine and 5% fetal bovine serum (FBS), at 37 ºC and 5% CO_2.**
    1. Seed initial frozen aliquots (2–5 x 10^5 cells) in a T-25 flask with 6 mL of RPMI 1640 media prepared in 2.1.
    2. Monitor cell density and viability.
        1. Take a representative aliquot of cells and stain them with an appropriate dye such as trypan blue.
        2. Measure the cell density using a hemocytometer or an automated cell counter.
    3. Before the cell reaches density of 1 x 10^6 cells/mL, spin the cells down at 100 x g for 5 min and transfer the cells to a T-75 flask with 25 mL of RPMI 1640 media prepared in 2.1. Grow the cells until cell density approaches 1 x 10^6 cells/mL.
    4. Transfer the cells to a T-150 flask in 38 mL of RPMI 1640 media (2.1) and grow the cells for an additional two days.
    NOTE: Lymphoblastoid cells grow in suspension. General guidelines for growing lymphoblastoid cells recommend cell densities to be kept between 0.2–1 x 10^6 cells/mL. Because of the large amount of material required in this technology, the present protocol is purposely written to grow cells to a cell density beyond what is commonly used. Cell growth protocols can be revised and adapted by the reader depending on the cell types to be used.
    5. Spin the cells down at 100 x g for 5 min, resuspend the cells in 10 mL of RPMI 1640 media (2.1) and pour the suspension into an 850 cm^2 roller bottle containing 500 mL of media. Incubate the cells under the same conditions as before but at this point start using constant rotation (~0.5 rpm).
    6. Monitor the cell growth and change the media when needed (Typically every 3–4 days). Let the cells grow to a density as close as possible to 2 x 10^6 cells/mL (1 x 10^9 cells in total).
    7. Once the desired cell count is reached, harvest the cells by spinning the cells down at 100 x g for 5 min and resuspend the cells in 36 mL of RPMI 1640 media (2.1). Transfer the cell suspension to a 50 mL conical tube. Take small aliquots for cell counting and DNA extraction.

2. **Crosslink by adding 1 mL of 37 % formaldehyde to reach a final concentration of 1%.**
CAUTION: Formaldehyde is a carcinogen and should be handled in a fume hood.
    1. Incubate the cells with rotation (~30–60 rpm) for 10 min at room temperature (RT).
    2. Quench the crosslink reaction by adding 2 mL of a 2.5 M glycine solution, for a final concentration of 125 mM.
    3. Incubate the cells with rotation (~30–60 rpm) for 5 min at RT. Pellet the cells by centrifuging at 100 x g for 5 min and wash the cells twice with ice cold 1x PBS.

3. Continue to next step or resuspend the pellet in 5 mL of Cell Lysis Buffer (5 mM Hepes, 85 mM KCl, 0.5% Igepal, protease inhibitors (PI)) and snap-freeze the suspension drop by drop in liquid nitrogen. Store the samples at -80 ºC.

NOTE: Prepare >25 mL of Cell Lysis Buffer without Igepal and PI and add those reagents fresh when ready to use. Avoid long-term storage of Cell Lysis Buffer once Igepal and PI have been added. Store Cell Lysis Buffer at 4 ºC for up to 3 months. Igepal is a nonionic, non-denaturing detergent. NP-40 can be potentially used as an alternative to Igepal but we have not used it in this context as NP-40 is not recommended when samples are eventually analyzed by mass spectrometry to characterize the DNA binding proteins.

3. Lysis and Shearing

1. Prepare >15 mL of Nuclei Lysis Buffer (10 mM EDTA, 50 mM Tris pH 8.0, 1 % SDS, PI).
   NOTE: Add PI fresh, just to the amounts to be used. Avoid long-term storage of Nuclei Lysis Buffer once PI has been added. Nuclei Lysis buffer can be stored at RT for up to 1 month.
2. Resuspend the pellet in 20 mL of Cell Lysis Buffer (containing Igepal and PI). Let the pellets thaw on ice and mix well once thawed. Let samples sit for an additional 10 min on ice. Centrifuge the cells at 400 x g at 4 ºC for 5 min.
3. Discard the supernatant and resuspend the cells in 6 mL of Nuclei Lysis Buffer with PI (volume can be increased if cells do not form a suspension).
4. **In preparation for sonication, divide the sample evenly in ~6–8 microfuge tubes (600 to 1,000 µL for each tube).**
   1. Place the samples on ice or a cold rack and sonicate the sample using a sonicator with a microtip. Use 65% amplitude with 5 x 20 s constant bursts, allowing the suspension to cool down for at least 40 s between pulses.
      NOTE: If volumes increase >10 mL per sample, a glass cooling cell can be used instead. More and longer bursts would be required then. Sonicators differ greatly. Conditions might have to be optimized by the reader. Other sonication alternatives can be used as well.
5. Centrifuge the cells at 12,000 x g for 10 min at 4 ºC, transfer the supernatant to a new tube, and discard the pellets.
6. Estimate total recovery using a fluorometric method.
7. Continue to next step or snap-freeze and store the sample at -80 ºC.
8. Optional: Take an aliquot, reverse crosslinks by incubating it overnight at 67 ºC in 0.3 M NaCl. Clean DNA and run it in Bioanalyzer to determine fragment lengths. The bulk of the fragments should be between 500 and 1,000 bp.
   NOTE: It is OK to stop at this point and store the samples at -80 ºC.

4. Hybridization Capture

1. Prepare 500 mL of 2x Hybridization (Hyb) Buffer (200 mM MES, 2M NaCl, 40 mM EDTA, 0.02 % Tween-20) and 500 mL of Wash Buffer (200 mM NaCl, 0.2 % SDS, 50 mM Tris pH 8). Store the Hyb and Wash Buffers at 4 ºC and room temperature, respectively, for up to 3 months. Use half of the 2x Hyb buffer to prepare a 1x Hybridization Buffer. Also store it at 4 ºC for up to 3 months.
   NOTE: The total volume of the hybridization capture will be twice the volume of the resulting sample in the previous step (as described so far, the sample volume would be around 6 mL and thus the total hybridization volume would be 12 mL).
2. Set aside beads equal to 5 % of the total hybridization capture volume (600 µL in this case). An appropriate and suitable magnet is needed to work with the beads in the following steps.
3. Wash the beads three times with 1x Hyb Buffer using twice the beads volume (1200 µL). Resuspend the beads in 1,200 µL of 2x Hyb Buffer. Place the tubes on the magnet until the solution clears and remove the buffer.
4. Perform pre-clearing reactions in 1.5 mL microfuge tubes with a maximum of 700 µL of sample per tube. Add enough beads (5%) and 2x Hyb Buffer to dilute the Hyb buffer to 1x (10 tubes with 600 µL of sample, 120 µL of washed beads resuspended in 2x Hyb Buffer, and 480 µL of 2x Hyb Buffer). Incubate them for 10 min at 31 ºC with end over end rotation (~30–60 rpm).
5. Place the tubes on a magnet until beads are immobilized and transfer the samples to new tubes. Discard the beads.
   NOTE: From this point forward, low-binding tubes are recommended.
6. Resuspend the capture oligonucleotides to a working solution of 10 pmol/µL and add 4 µL of this solution to each tube.
   NOTE: Different regions can act as control for each other, but it is recommended to include a capture oligonucleotide with its sequence scrambled to serve as a true negative control.
7. Incubate the tubes for 40 min at 42 ºC with end over end rotation (~30–60 rpm).
8. During the incubation, set aside the beads needed (0.3 µL/pmol of capture oligonucleotide). Wash the beads three times as described in step 4.3, but resuspend the cells in 1 x Hyb Buffer.
9. Add the washed beads to the sample and incubate the samples for 30 min at RT with end over end rotation (~30–60 rpm). Place the tubes on the magnet and remove the supernatant once the beads have been immobilized.
10. **Washing the beads**
    1. Add 1.2 mL of Wash Buffer and incubate the sample for 5 min with end over end rotation (~30–60 rpm) at RT. Place tubes on the magnet and wait a couple of min until solution clears. Remove the buffer and repeat this step.
    2. Add 1.2 mL of Wash Buffer and incubate it for 1 h with end over end rotation (~30-60 rpm) at RT. Place the tubes on the magnet and wait a couple of min until solution clears. Remove the buffer and repeat this step reducing the incubation time to 30 min the second time.
    3. Add 200 µL of Wash Buffer and incubate it for 15 min with end over end rotation (~30–60 rpm) at 31 ºC. Place the tubes on the magnet and wait a couple of min until solution clears. Remove the buffer.
    4. Add 200 µL of PBS and incubate it for 5 min with end over end rotation (~30–60 rpm) at RT. Place the tubes on the magnet and wait a couple of min until solution clears. Remove the buffer and repeat this step.
11. **Simple elution**
    1. Add 40 µL of molecular grade water to the beads and incubate it at 94 ºC for 5 min.
    2. Place the tubes on the magnet and wait a couple of minutes until the solution clears.
    3. Transfer the supernatant into a new tube. Discard the beads.
    4. Proceed to the proteomic analysis or store the sample at -80 ºC.

NOTE: This type of elution works well for later qPCR analysis and some proteomics measurements, but in other instances two alternative approaches described below yield "cleaner" eluates.

12. **DNase elution.**
    1. Add 40 µL of 1x DNase Buffer and 2 U of DNase to each tube.
    2. Incubate the tubes for 15 min at 37 ºC.
    3. Place the tubes on a magnet and wait a couple of min until the solution clears.
    4. Remove the eluted samples from the beads and place the samples in a new low-binding tube.
    5. Incubate the tube at 94 ºC for 7 min to inactivate the enzyme and proceed to proteomic analyses.
       NOTE: Lower DNase concentrations can be used if interference with proteomic analyses is observed. It will not be possible to calculate the capture enrichments by qPCR if this elution method is used since the DNase treatment completely destroys the DNA material in the sample.
    6. Proceed to proteomic analysis or store the sample at -80 ºC.

13. **Toehold elution[17].**
    NOTE: This method is not recommended for longer oligonucleotides (>40 bases). This method requires for the capture oligonucleotides to contain an additional 8 bases that do not overlap with the genomic sequence and complementary release oligonucleotides in order to displace the target from the capture oligonucleotides.
    1. Dilute 2 nanomoles of release oligonucleotides in 40 µL of SSC buffer.
    2. Incubate it for 20 min at RT.
    3. Place the tubes on a magnet and wait a couple of minutes until the solution clears.
    4. Remove the solution from the beads.
    5. Proceed to proteomic analysis or store the sample at -80 ºC.

5. Evaluation of Capture Yield by qPCR with Primers for the Target Capture Region of Interest

1. Use an appropriate software or online tool to design primers and probes for the target region.
   NOTE: In this experiment, the promoter region of the *DUSP3* and *DUSP5* were targeted. The primers and probes used are the following. *DUSP3*: Primer 1 - GTG TTT AGG TTC CCT GAT CC, primer 2 - CGG AAT GTC TGC CTT CG, FAM-labeled probe - TCC ATG CCC GAA TTG TGA CGT CGG; *DUSP5*: Primer 1 - TGA GAA AGC CCT GAT GTG TC, primer 2 - GCC TTC AGA AAC CCC AAA AG, FAM-labeled probe TGC ATT TAG AGC AGC CAG ATG AGG TT.
2. Take an aliquot of the eluate (no more than 10 µL), and dilute 1:10 with molecular grade H_2O.
3. Extract the DNA of one of the aliquots taken in step 2.1.7 and use that DNA sample to make a standard curve by doing five serial 1:10 dilutions in molecular grade H_2O.
4. Prepare a qPCR master mix (wide variety of commercial alternatives), add primers and probes and dispense appropriate amount (15 µL if running a 20 µL reaction) in each well.
5. Add 5 µL of either: a) diluted sample, b) the standard curve or c) molecular grade H_2O (to serve as no template control) to each well.
6. Seal the plate and centrifuge it at 100 x g for 1 min.
7. Run in a qPCR system with the following parameters: 5 min at 95 ºC followed by 40 cycles of 95 ºC for 20 s, 59 ºC for 20 s and 72 ºC for 25 s.
8. Use the standard curve to asses yields and scrambled capture to assess capture specificity.

Representative Results

Due to the need for large input amounts of chromatin for HyCCAPP to succeed, cells are grown to relatively high levels of confluency. Trypan blue staining is used to confirm that cell death rates are moderate (<10 %). In single copy experiments, chromatin content prior to hybridization capture needs to be in the femtomolar range, which usually requires at least 10^9 cells as starting material. Prior to full scale experiments, it is recommended to test capture oligonucleotide performance in smaller batches. Hybridization captures using a titration of oligonucleotides are shown in **Figure 1**. The number of capture oligonucleotides is gradually increased while the total concentration remains constant. This experiment clearly shows three key aspects. First, it helps to determine, for that particular region, how many oligonucleotides are needed to reach optimal (and maximal) hybridization efficiencies. Secondly, it reveals if there are any obvious detrimental interactions between a particular set of oligonucleotides. Lastly, it shows if there are any oligonucleotides with poor specificity that carry substantial amounts of background (in this case measured by qPCR analysis with primers targeting other regions in the genome).

The capture hybridization yields will appear modest due to the cross-linked nature of the chromatin sample. A large proportion of fragments in the crosslinked chromatin will not be amenable for hybridization capture. Serial hybridization experiments demonstrate this (**Figure 2**). If the hybridization capture is successfully run, a subsequent capture experiment of the same chromatin material using a different set of capture oligonucleotides will result in comparable (slightly lower) yields than when using fresh chromatin (~11% lower on average). But if the same region is targeted again with the same set of oligonucleotides in a second capture experiment using the chromatin sample remaining after the first capture experiment, the yield will decrease about 90%, confirming that most of the hybridization-amenable material has been captured. If such a decrease in capture is not observed after subsequent captures with the same set of oligonucleotides, it indicates a technical problem in the hybridization capture step.

In simple systems like yeast, we have observed capture efficiencies approaching 4%[5], which led to the identification of 9 proteins differentially enriched in 2 regions when yeast was grown under different conditions. In human cells however, having a genome ~250 times larger than yeast, hybridization yields are typically below 1% (**Figure 3**). As **Figure 3** shows, a single experiment might yield insufficient amounts for proteomics analyses. In that case, several captured samples will have to be pooled in order to perform a subsequent mass spectrometry

analysis. Alternatively, lower concentrations of captured target region chromatin can be used for targeted analysis approaches such as western blotting and mass spectrometry-based selected reaction monitoring assays.

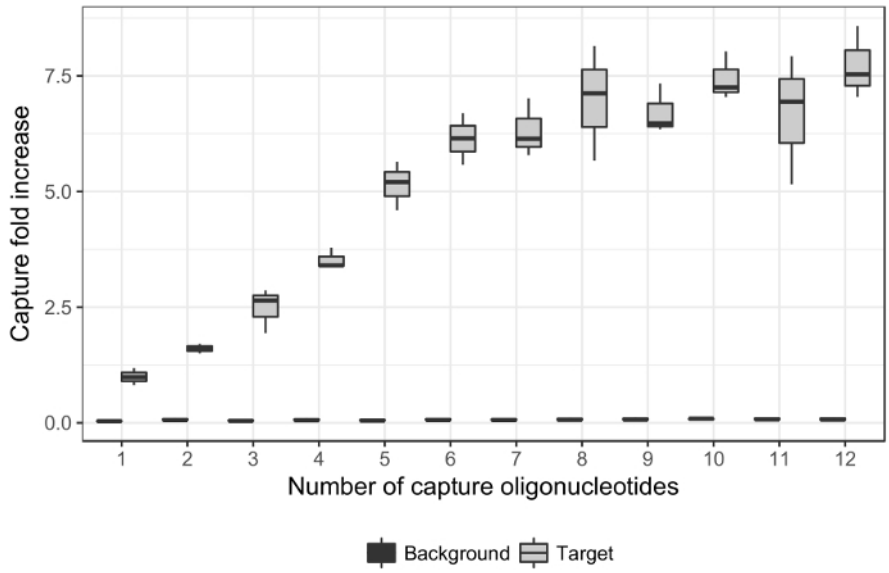

Figure 1: Box plot representation of capture oligonucleotide titration. Increasing number of capture oligonucleotides were tested and analyzed by qPCR. Background refers to qPCR assays targeting other regions in the genome. Values are presented as fold change from hybridization capture using a single oligonucleotide. Whiskers represent the minimum and maximum values. Please click here to view a larger version of this figure.

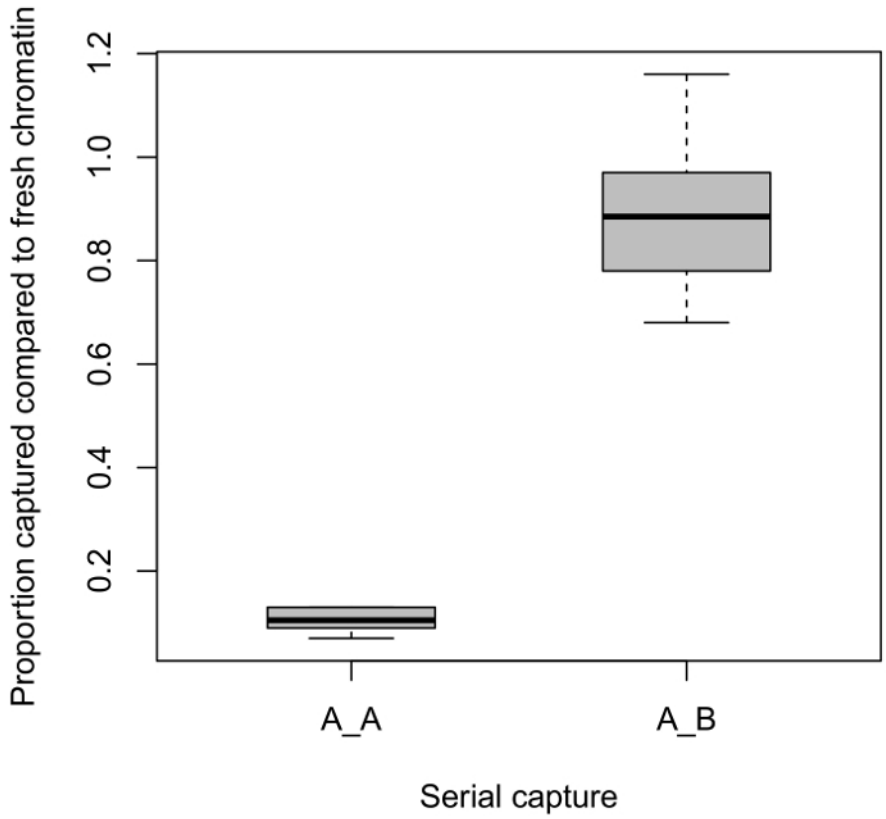

Figure 2: Box plot representation of serial captures. Sequential hybridization captures with the same set of oligonucleotides (A_A) show a strong drop in enrichment compared to sequential hybridization captures using different set of oligonucleotides (A_B) ($p = 3.88\ e^{-5}$). Enrichment is measured by qPCR and values are shown relative to captures using fresh chromatin. Whiskers represent the minimum and maximum values. Please click here to view a larger version of this figure.

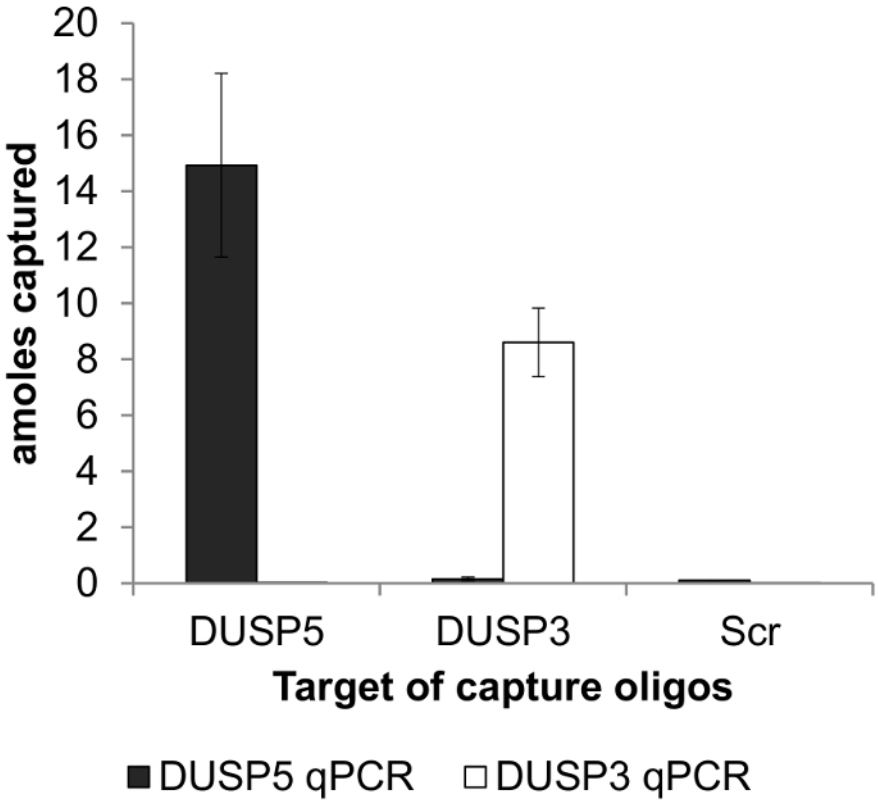

Figure 3: Hybridization capture using human cells. Hybridization capture of regions located in different chromosomes. Each region serves as a negative control for the other. Additionally, a hybridization capture is performed using the scrambled (Scr) sequence as a true negative control. Error bars represent standard deviation. Please click here to view a larger version of this figure.

Discussion

The HyCCAPP method described here has many unique features that make it a powerful approach to uncover DNA-interactions that otherwise would remain elusive. The nature of the process gives HyCCAPP the flexibility to work in different organisms and regions of the genome. It is a method, however, that has several limitations to be considered.

HyCCAPP is a method that avoids any cell modifications so that it can potentially be applied in primary cells, cell culture systems, or even tissue samples. Because of that, however, it requires samples to be crosslinked, which reduces sensitivity in the proteomic analysis and therefore requires large input amounts. There are emerging approaches that rely on CRISPR and biotinylation of surrounding proteins that can function without crosslinking reactions[12,13,14,18]. These approaches can be very powerful if plasmid insertions do not represent a constraint, but can only be used in cell culture systems. Other approaches are very well suited to identify general protein associations based on a chromatin state or on the presence of a particular transcription factors, but are not meant to study individual loci[9,10,19].

The HyCCAPP approach presents an alternative approach with broad applicability, but it is likely that applications in different cell systems or organisms will require optimization. The most critical optimization step in the approach involves the design and selection of capture oligonucleotides. A series of small scale tests should show if the designed oligonucleotides are appropriate and if the sequence in the region is effectively captured. There are several factors that should be considered when designing and testing capture oligonucleotides, such as oligonucleotide length, degree of specificity to the target region, and interactions among the set of oligonucleotides to be used. All these interactions can be readily tested by using a qPCR assay specific to the target region. If desired, a ChIP-Seq-like workflow can be used to sequence the resulting DNA fragments and assure that the specificity of the enrichment is satisfactory across the entire genome[5].

Finally, the cross-linking conditions are also highly important in the HyCCAPP procedure, and we have observed noticeable differences between different systems, obtaining the best results in human cells by cross-linking with 1% formaldehyde, whereas in yeast, cross-linking with 3% formaldehyde yielded more reproducible results. It is possible that the more intense cross-linking with 3% formaldehyde in human cells with a more complex chromatin structure does not provide sufficient accessibility for hybridization capture oligonucleotides. It is also conceivable that the presence of a cell wall in yeast hinders formaldehyde access into the cells, increasing the amount of formaldehyde needed for successful crosslinking levels in the chromatin.

Performing ChIP and other capture approaches targeting crosslinked human chromatin, we have observed that only about 1% of the target DNA can actually be captured. For most DNA-based approaches and analyses, this is not a limiting factor. However, unlike ChIP, where the final readout is based on amplifiable DNA, HyCCAPP relies on protein content for the final readout. Due to this intrinsic limitation, large amounts of input material (cultured cells in the experiments presented here) are required: this constrain should be carefully considered before exploring this methodology, especially when applied to single-copy regions. Not all systems will be able to produce the number of cells needed, or the cost of

growing the required cell numbers might be prohibitive. The input amounts that HyCCAPP requires (~10^9 cells) is comparable to other methods that rely on crosslinked material with input amounts ranging between 10^8 and 10^{11} cells[9,11,15]. Future modifications of the HyCCAPP technology will explore approaches to artificially increase copy numbers of the target regions, to make this method more broadly applicable. At the same time, we will work to continue increasing the overall efficiency of the process that together with continuous advancements in mass spectrometry should reduce the input amounts needed and make this technology feasible in more systems.

Biotinylation based approaches using CRISPR, like enCHiP[18] and others[13,14], have shown to be very effective by eliminating the need of crosslinked material, increasing yields, and significantly reducing input requirements. The elaborate processing of cells in these methods however, does not allow these techniques to be applied to tissue samples, a direction that we have begun to pursue with HyCCAPP, and that is producing promising results.

Disclosures

There are no conflicts of interest to disclose.

Acknowledgements

This work was supported by NIH Grants P50HG004952 and R01GM109099 to MO.

References

1. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature.* **489** (7414), 57-74 (2013).
2. Galas, D. J. and Schmitz, A. DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* **5** (9), 3157-70 (1978).
3. Simon, J. M., Giresi, P. G., Davis, I. J. and Lieb, J. D. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat. Protoc.* **7** (2), 256-267 (2012).
4. Dai, Y., Kennedy-Darling, J., Shortreed, M. R., Scalf, M., Gasch, A. P. and Smith, L. M. Multiplexed Sequence-Specific Capture of Chromatin and Mass Spectrometric Discovery of Associated Proteins. *Anal. Chem.* **89** (15), 7841-7846 (2017).
5. Guillen-Ahlers, H. *et al.* HyCCAPP as a tool to characterize promoter DNA-protein interactions in Saccharomyces cerevisiae. *Genomics.* **107** (6) (2016).
6. Kennedy-Darling, J. *et al.* Discovery of Chromatin-Associated Proteins via Sequence-Specific Capture and Mass Spectrometric Protein Identification in Saccharomyces cerevisiae. *J Proteome Res.* **13** (8), 3810-3825 (2014).
7. Guillen-Ahlers, H., Shortreed, M. R. R., Smith, L. M. M. and Olivier, M. Advanced methods for the analysis of chromatin-associated proteins. *Physiol Genomics.* **46** (13), 441-447 (2014).
8. Lambert, J.-P., Fillingham, J., Siahbazi, M., Greenblatt, J., Baetz, K. and Figeys, D. Defining the budding yeast chromatin-associated interactome. *Mol Syst Biol.* **6**, 448 (2010).
9. Soldi, M. and Bonaldi, T. The Proteomic Investigation of Chromatin Functional Domains Reveals Novel Synergisms among Distinct Heterochromatin Components. *Mol Cell Proteomics.* **12** (3), 764-780 (2013).
10. Wang, C. I. *et al.* Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. *Nat Struct Mol Biol.* **20** (2), 202-209 (2013).
11. Byrum, S. D., Raman, A., Taverna, S. D. and Tackett, A. J. ChAP-MS: A Method for Identification of Proteins and Histone Posttranslational Modifications at a Single Genomic Locus. *Cell Rep.* **2** (1), 198-205 (2012).
12. Fujita, T. and Fujii, H. Efficient isolation of specific genomic regions and identification of associated immunoprecipitation (eChIP) using CRISPR. *Biochem. Bioph. Res. Co.* **439** (1), 132-136 (2013).
13. Liu, X. *et al.* In Situ Capture of Chromatin Interactions by Biotinylated dCas9. *Cell.* **170** (5), 1028-1043.e19 (2017).
14. Myers, S. A., Wright, J., Zhang, F. and Carr, S. A. CRISPR/Cas9-APEX-mediated proximity labeling enables discovery of proteins associated with a predefined genomic locus in living cells. *bioRxiv.* at <http://biorxiv.org/content/early/2017/07/04/159517.abstract> (2017).
15. Déjardin, J. and Kingston, R. E. Purification of Proteins Associated with Specific Genomic Loci. *Cell.* **136** (1), 175-186 (2009).
16. Buxton, K. E. *et al.* Elucidating Protein-DNA Interactions in Human Alphoid Chromatin via Hybridization Capture and Mass Spectrometry. *J. Proteome Res.* **16**(9), 3433-3442 (2017).
17. Kennedy-Darling, J., Holden, M.T., Shortreed, M.R., & Smith, L.M. Multiplexed programmable release of captured DNA. *Chembiochem.* **15** (16), 2353-2356 (2014).
18. Fujita, T. and Fujii, H. Isolation of specific genomic regions and identification of associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Chromatin Protoc. Third Ed.*, 43-52 (2015).
19. Lambert, J.-P., Mitchell, L., Rudner, A., Baetz, K. and Figeys, D. A Novel Proteomics Approach for the Discovery of Chromatin-associated Protein Networks. *Mol Cell Proteomics.* **8** (4), 870-882 (2009).

