Matthew J. Oberley, Peggy J. Farnham[1]

McArdle Laboratory for Cancer Research

University of Wisconsin Medical School

1400 University Ave

Madison, WI 53706

[1]corresponding author: 608-262-2071 (tel); 608-262-2824 (fax);

farnham@oncology.wisc.edu

INTRODUCTION

Many methods have been developed to identify genomic targets of DNA-binding factors. Here we outline and discuss our adaptation of the chromatin immunoprecipitation (ChIP) assay to a high-throughput microarray based method for discovering genomic regions occupied by human DNA-binding proteins. Others, primarily using yeast model systems, have also explored the use of coupled chromatin immunoprecipitation and microarrays to identify large sets of binding sites of DNA-binding proteins. For example, binding sites for yeast transcriptional regulators such as SBF, MBF, Gal4, Ste12 and Rap1 have been identified. [1, 2, 3] Other studies have focused on chromatin remodeling factors, such as components of the RSC complex, [4, 5] and

components of the DNA replication machinery, such as ORC and MCMs.[6] In certain cases, the investigators have monitored all intergenic regions and/or known promoters in the yeast genome.[1, 3] In other cases, the authors have taken advantage of the relatively small size of the yeast genome to spot probes which span the entire genome (both intergenic regions and open reading frames) onto a microarray.[2, 6] A recent, Herculean effort by Lee et al.[7] has utilized ChIPs and microarrays to identify the global binding locations of the majority of known yeast transcription factors and used the resulting data along with data from mRNA expression array analysis to define the regulatory networks that exist in yeast.

Because the human genome (3.2 Gb) is three orders of magnitude larger than the yeast genome (12 Mb), a chip containing all human intergenic regions and open reading frames has not yet been created. The largest effort to date consists of a set of oligonucleotide arrays that span the smaller chromosomes 21 and 22 at 35 bp intervals.[8] However, this array has not to date been used to identify sites for DNA binding factors, rather it is been used to assess the sum transcriptional activity of these chromosomes. A microarray, more limited in scope but allowing a more detailed study of specific regions, which includes the sequence spanning from –700 to +200 of 1444 human genes has been utilized to find novel promoters bound by E2F family members.[9] Although a step in the right direction, such an array by necessity is biased in that the particular 1444 promoters analyzed were chosen by a set of criteria. To eliminate such bias, our laboratory has utilized a microarray that contains 7,776 CpG island clones to identify genomic sites to which E2F and pRb family members are recruited.[10, 11] Of particular interest to the study of transcription factors is that CpG islands tend to be found in intergenic regions and at

the 5' ends of genes. Thus, use of CpG arrays now allows an unbiased analysis of many thousands of human intergenic regions.

CpG islands, found in or near approximately 50% of human promoters, are identified by three primary characteristics: they are more than 200bp long, have over 50% GC composition, and retain an observed/expected ratio of CpG dinucleotides greater than 0.6.[12] The CpG library used on the microarrays was originally prepared in the laboratory of Bird[13] and consists of approximately 60,000 independent clones which represent CpG islands which were not hypermethylated in human male genomic DNA. Clearly one limitation of this array is that any CpG island which was hypermethylated in the starting genomic DNA would not be represented in the library. However, the technique could be modified to allow the collection of all CpG islands (Figure 1). Alternatively, new libraries could be prepared from different starting tissues, allowing the creation of arrays containing CpG islands that are hypo- vs. hypermethylated in specific tissues, in tumors, or in different developmental stages. Approximately 8000 clones representing hypomethylated CpG islands were arrayed by the Huang lab[14] onto slides for high-throughput study of DNA methylation changes in human cancer. As described below, we have now developed protocols which allow these arrays to be utilized to perform a large-scale identification of genomic regions occupied by DNA-binding proteins under physiologically relevant, in vivo conditions. As illustrated in Figure 2, the assay subsumes three major areas: i.) chromatin immunoprecipitation, ii.) DNA labeling and hybridization, and iii.) data analysis. Therefore, this chapter has been divided into these three major areas; each section begins with the protocol and ends with Critical Points and Reagents; illustrative figures are also provided.

I. CHROMATIN IMMUNOPRECIPITATION

We have used the following protocol to successfully immunoprecipitate sequences specifically associated with E2F family members, pocket proteins (Rb, p107, p130), β-catenin and TCF family members, RNA polymerase, Myc family members, histones, etc. Therefore, we believe that this protocol is applicable to a wide variety of different types of DNA binding factors. We have utilized many different cell types (epithelial, fibroblast, blood cells, etc) and have also adapted it for use with mouse and human tissues. Although minor modifications may be required to develop the optimal set of conditions for a particular antibody and/or cell type, we anticipate that the following protocol will provide a good starting point for most transcription factors. For further information on the ChIPs assay see Weinmann and Farnham.[15]

Day 0

Block Staph A cells: The preparation of Staph A cells is described below in the Reagents section. Thaw 100 ul of prepared Staph A for approximately every 5 x 10^7 cells that you begin with. Add 10 ul of herring sperm DNA (10 mg/ml) and 10 ul of BSA (10 mg/ml) for every 100 ul of Staph A. Incubate the mixture on the rotating platform at 4°C overnight then, on the following day, microfuge the mixture for 3 min. Remove the supernatant and wash the Staph A pellet twice with 1X dialysis buffer. Finally, resuspend the Staph A cells in a volume of 1X dialysis buffer equal to the original starting volume.

Day 1[*]

! 1. Add formaldehyde directly to tissue culture media to a final concentration of 1%. Incubate adherent cells on a shaking platform and suspension cells on a stir plate for 10 min at room temperature. We use between 2 X 10^6 and 1 x 10^7 cells per antibody per timepoint for standard chromatin immunoprecipitation assays depending on the antisera.

2. Stop the crosslinking reaction by adding glycine to a final concentration of 0.125 M. Continue to rock or spin the cells at room temperature for 5 min.

3. For adherent cells, pour off the media and rinse the plates twice with cold 1X PBS. For suspension cells, centrifuge at 1500 rpm and wash the cell pellet twice with cold 1X PBS (for suspension cells then proceed to step 6).

4. For adherent cells, add an appropriate volume (try 5 mls per 500 cm2 dish) of 1X PBS or a 0.0025% trypsin solution (a 1:5 dilution of a 1X Trypsin-EDTA tissue culture grade solution diluted in 1X PBS). Incubate the cells at 37°C for 10 min if using trypsin. The trypsin step is useful for cells which are difficult to swell; for cell types that are easily swelled, this step may not be necessary.

! 5. Following addition of trypsin or 1X PBS, scrape adherent cells from dishes. If using trypsin to help swell the cells, inactivate the trypsin by adding a small amount of serum. Centrifuge the scraped adherent cells at 1500 rpm and wash the cell pellet once with 1X PBS plus PMSF (10 ul of the stock solution per ml).

---

[*] If a step in the protocol is marked with an exclamation point, that indicates that the reader is referred to a Critical Point, located immediately following that section of the protocol.

6. Resuspend the cell pellet in cell lysis buffer plus the protease inhibitors PMSF (10 ul of the stock solution per ml), aprotinin (1 ul of the stock solution per ml) and leupeptin (1 ul of the stock solution per ml). The final volume of cell lysis buffer should be sufficient so that there are no clumps of cells. Incubate the cells on ice for 10 min. Cells can also be homogenized on ice with a dounce homogenizer (B pestle) several times to aid in nuclei release. The cell lysis conditions, buffers, and homogenizing may need to be optimized for nuclei isolation from individual cell types.

7. Centrifuge the homogenate at 5,000 rpm for 5 min at 4°C to pellet the nuclei.

8. Discard the supernatant and resuspend the nuclei pellet in 1 ml nuclei lysis buffer plus protease inhibitors. Incubate on ice for 10 min. If the nuclei are too dense, resuspend them in a larger volume of nuclei lysis buffer.

9. Sonicate the chromatin to an average length of approximately 0.5-1 kb, and make sure to keep the samples cold during the sonication procedure. The sonication time and number of pulses will vary depending on the sonicator used, the cell type, and the extent of crosslinking. As a starting point, we have performed the sonication step using 4 pulses of 15 seconds each at setting 3 on a Fisher model 60 sonic dismembrator. In between pulses, samples should be allowed to cool for 30 seconds on ice. To aid in sonication, it may also be helpful to add 0.1 grams of glass beads (212-300 microns; Sigma, G-1277) prior to sonication. After sonication, microcentrifuge samples at 14,000 rpm for 10 min at 4°C. Carefully remove the supernatant and transfer to a new tube. At this point, the chromatin can be snap frozen in liquid nitrogen and stored at –80°C.

10. Preclear the chromatin by adding 10 ul of preblocked Staph A cells for every 10^7 crosslinked cells.

11. Incubate the chromatin/Staph A mixture on a rotating platform at 4°C for 15 min, then microcentrifuge the mixture at 14,000 rpm for 4 min.

! 12. Transfer the supernatant to a new tube and divide equally among IP samples. For example, if starting with 1 x 10^8 cells, then divide into 10 samples; 8 to be precipitated with the antibody specific to the DNA binding protein of interest, 1 for an IgG control, and 1 to save as total reference DNA. Adjust the final volume of each sample with 2 times the chromatin volume of IP dilution buffer (plus protease inhibitors). The volume of each sample should be between 200 ul and 600 ul. We generally add 1 ug of the specific antibody to the appropriate samples. However, the optimal antibody concentration may vary for individual antibodies; initial tests can be performed with 0.5, 1, 2, and 5 ug of antibody per sample.

13. Incubate the chromatin samples with the antibodies on a rotating platform at 4°C overnight (or for at least 3 hours).

Day 2

14. If you are not using rabbit polyclonal antibodies (i.e. you are using a mouse monoclonal antibody, a goat polyclonal antibody, etc.), add 1 ug of an appropriate secondary antibody and incubate for an additional 1 hour. We have found that rabbit polyclonal antibodies associate well with Staph A cells; therefore secondary antibodies should be from a rabbit.

15. Add 1 ul of blocked Staph A cells per $1e^6$ cells to each sample, and incubate them on a rotating platform at room temperature for 15 min.

16. Microcentrifuge samples at 14,000 rpm for 4 min. Discard the supernatant.

17. Wash pellets twice with 1.4 ml of 1X dialysis buffer (*if you are using a monoclonal antibody, omit the sarkosyl from the buffer) and four times with 1.4 ml of IP wash buffer (*pH 8.0 for monoclonal antibodies). For each wash, dissolve the pellet in 700 ul of buffer then add an additional 700 ul of buffer and incubate samples on a rotating platform for 3 min. Next, microcentrifuge samples at 14,000 rpm for 4 min. Try to remove as much buffer as possible after each wash without aspirating the Staph A cells.

18. After the final wash, remove as much wash buffer as possible, and dry spin the tubes for 4 min at 4°C. Then remove all traces of wash buffer.

19. Add 150 ul of IP elution buffer, and vortex on setting 3 for 15 min at room temp. Spin the tube for 4 min, and transfer the elution buffer to a fresh tube.

20. Add an additional 150 ul of IP elution buffer to the Staph A pellet and vortex again on setting 3 for 15 min at room temp. Spin the tube for 4 min, remove the supernatant and combine the two elution volumes.

21. After the second elution, microcentrifuge the samples at 14,000 rpm for 4 min to remove any traces of Staph A cells. Transfer the supernatants to new tubes. Add 12 ul of 5M NaCl to a final concentration of 0.3 M. Remember to include the "total reference DNA" sample at this point to reverse crosslinks. For the total sample, use only 20% of starting IP volume, adjusting to 300 ul with IP elution buffer. Incubate the samples at 67°C for 4-5 hours to reverse the formaldehyde crosslinks. After the 4-5 hour incubation, add 2.5 volumes of ethanol to each sample and precipitate them at –20°C overnight.

Day 3

27. Microcentrifuge the samples at 14,000 rpm for 15-20 min at 4°C. Discard the supernatant, and re-spin to remove any residual ethanol. Allow the pellets to air dry completely.

28. Dissolve each pellet in 100 ul of TE. Add 25 ul of 5X PK buffer and 1.5 ul of proteinase K (25 mg/ml) to each sample. Incubate the samples at 45°C for 1-2 hours.

30. To remove all protein and contaminants from the immunoprecipitated chromatin we purify the chromatin with a Qiaquick PCR cleanup kit (Qiagen, 28106). We follow the Qiagen protocol except that we elute by adding 30ul of buffer EB to the column and allow to stand 1 min at room temperature. The samples are then centrifuged at maximum speed for 1 min in a microcentrifuge.

! 31. At this point, the success of the chromatin immunoprecipitation portion of the experiment can be monitored by a standard PCR reaction using primers specific for a region of the genome thought to be bound by the factor of interest.

CRITICAL POINTS:

! (Step 1) For attached cells, we have noticed that it is critical to formaldehyde crosslink cells on the dish before trypsinizing, as we have observed that DNA-binding factors can relocalize after trypsinization. This precludes an experimental design that incorporates live cell-sorting after trypsinization to create relatively pure populations of cells. However, we have found that nuclei prepared from crosslinked cells can be easily sorted

for DNA content and subsequently immunoprecipitated with antibodies specific to chromatin binding proteins.

! (Step 5) At this point, the cell pellets can be snap-frozen and stored at –80°C. This is helpful if transcription factor binding in different cell populations (which are not necessarily ready at the same time) are to be compared in the chromatin immunoprecipitation assay.

! (Step 31) Do not proceed with the hybridization portion of the experiment unless the signal obtained in a test PCR experiment shows higher signal in the antibody samples than in the IgG control. Care must be taken to avoid cross-contamination of samples.

REAGENTS

Cell Lysis buffer

5 mM PIPES pH 8.0

85 mM KCL

0.5% NP40

add PMSF just before use

Nuclei Lysis buffer

50 mM Tris-Cl pH 8.1

10 mM EDTA

1% SDS

add PMSF, aprotinin, and leupeptin just before use

IP Dilution buffer

0.01% SDS

1.1% Triton X 100

1.2 mM EDTA

16.7 mM Tris-Cl pH 8.1

167 mM NaCl

add PMSF, aprotinin, and leupeptin just before use

1X Dialysis buffer

2 mM EDTA

50 mM Tris-Cl pH 8.0

0.2 % Sarkosyl (omit for monoclonal antibodies)

add PMSF just before use

IP Wash buffer

100 mM Tris-Cl pH 9.0 (8.0 for monoclonal antibodies)

500 mM LiCl

1% NP40

1% deoxycholic acid

add PMSF just before use

Elution buffer

50 mM NaHCO3

1% SDS

5X PK buffer

50 mM Tris-Cl pH 7.5

25 mM EDTA

1.25% SDS

Protease Inhibitors

100 mM phenylmethylsulfonyl fluoride (PMSF; Sigma, P-7626) in isopropanol, use at 1:100 dilution.

10 mg per ml aprotinin (Sigma, A-1153) in 0.01 M HEPES pH 8.0, use at 1:1,000 dilution.

10 mg per ml leupeptin (Sigma, L-2884) in water, use at 1:1,000 dilution.

Staph A Cells

Resuspend 1 gram of lyophilized protein A-positive staphylococcus aureus cells, whole cells (Cowan 1 strain), heat-killed, fixed in formalin, (Calbiochem, 507862**)** in 10 ml of 1X dialysis buffer. Centrifuge at 10,000 rpm for 5 min at 4°C. Discard supernatant and repeat wash. Resuspend in 3 mls of 1X PBS plus 3% SDS and 10% BME. Boil for 30 min. Centrifuge at 10,000 rpm for 5 min. Wash in 1X dialysis buffer and centrifuge at 10,000 rpm for 5 min. Repeat wash. Resuspend in 4 mls of 1X dialysis buffer. Divide into 100 ul aliquots, snap freeze and store in liquid nitrogen.

II. GENERATING CHROMATIN AMPLICONS

In our previously published studies, we began by pooling between 28 and 50 individual ChIP assays (using 1 x 10^7 cells/assay) to obtain enough DNA to probe replicate arrays. Although tedious, it is generally possible to obtain this requisite number of cells using immortalized human cell lines. However, if multiple experimental points are to be analyzed or if primary human cells or tissues are to be used, collecting this large number of cells becomes onerous. Therefore, we have adapted a ligation-mediated PCR technique (LMPCR) which is based on the method of Ren et al.[13] to create amplicons of individual IPs. We have found that as little as one-half of one ChIP assay can be efficiently amplified to give amplicons that accurately represent the starting population . The ligation-mediated PCR technique simply involves blunt ending the chromatin, ligating a unidirectional double-stranded oligonucleotide linker, and PCR amplifying the resultant DNA population. Other groups have utilized random priming of the starting chromatin material with comparable results.[12] Recently, a promising amplification method has been developed that utilizes in-vitro transcription of the chromatin templates to create RNA amplicons.[26]. Because this procedure results in a linear rather than exponential amplification, this method has the potential to reduce potential bias introduced by PCR-based amplification schemes. Indeed, this form of sample amplification is commonly used in microarray studies of gene-expression where mRNA samples are limited.[27]

To ensure that the LMPCR reaction has not resulted in a loss of difference between the experimental and control samples, we amplify the IP samples as well as an identical volume of the total reference DNA (0.2% of the starting IP). This amplified input DNA

will be used as a reference on the CpG microarray. We also commonly amplify a non-specific IgG ChIP control reaction to demonstrate that enrichment is due to the specific antibody-epitope interaction. Therefore, three different samples can be prepared which can be used in two different hybridizations comparisons (i.e. IP v. input, and IgG v. input). Please note that this LMPCR amplification step is a new addition to our protocol; to call attention to this fact, both Day 1 and Day 2 have been identified with an exclamation point; all steps in these sections are new.

! Day 1

1. The two unidirectional linkers oligoJW102 (5' gcggtgacccgggagatctgaattc 3') and oligoJW103 (5' gaattcagatc 3') are annealed by combining 6.7 ul of 100 uM oligoJW102, 6.7 ul of 100 uM oligoJW103, and 86.6 ul H_20. This mixture is boiled for 5 min in a water bath, then allowed to slowly cool to room temperature. The annealed linker can be stored indefinitely at -20^{o}C.

2. Because sonication of chromatin creates overhanging ends, the chromatin is blunt ended by mixing the chromatin sample derived from an experimental sample, a IgG sample, or 10 ng of input with the following: 11 ul T4 DNA Polymerase Buffer, 5 ul 10X BSA, 5 ul 2 mM dNTPs, 1 ul T4 DNA Polymerase (New England Biolabs, M0203S), and H_2O to 110 ul. This mixture is placed at 37^{o}C for 45-60 min, and then each sample is purified with a Qiaquick PCR purification kit, according to manufacture's protocol, eluting with 30 ul of elution buffer.

3. Next, double stranded unidirectional linker from step one is ligated to the blunted chromatin by adding 27 ul of the blunted chromatin, 10.3 ul H_20, 5 ul 10X T4

DNA Ligase Buffer, 6.7 ul Annealed JW102/JW103 linker, and 1.0 ul T4 DNA ligase (New England Biolabs, M0202S). This mixture is placed at 16°C overnight.

! Day 2

4. The following morning, each sample is purified with the Qiaquick PCR purification kit, as above, eluting with 30 ul of elution buffer.

5. The linker-ligated chromatin is then PCR amplified to create stocks of amplicons from which virtually unlimited amounts of DNA can be generated for use with the CpG-island microarrays. Each PCR reaction contains 5 ul 10X Taq polymerase buffer, 7 ul 2 mM dNTPs, 3 ul $MgCl_{2,}$, 6.5 ul betaine, 2.5 ul oligoJW102 (20 uM), 1 ul Taq (Promega, M1861), and 25 ul of the linker-ligated chromatin.

6. PCR is performed using the following protocol:

55°C for 2 min
72°C for 5 min
95°C for 2 min
] 1 cycle

95°C for 0.5 min
55°C for 0.5 min
72°C for 1 min
] 15 cycles

72°C for 4 min
4°C for ∞

7. Following the PCR, each reaction is purified using the Qiaquick PCR purification kit according to the manufacturer's instructions, eluting in 30 ul of elution buffer.

8. The PCR protocol (Steps 5, 6, and 7) is repeated until enough amplicon is made for a master stock (1-10 ug), which can be stored at -20^{o}C. After several rounds of amplification, 5 ul of the amplicon should be visible on a 1% agarose gel, with fragments

ranging from 200-600 bp, even if the starting chromatin was sheared to a much larger size. We always amplify all samples to be compared using the same number of overall cycles to avoid introducing bias between the IP and total sample. The number of times required to repeat the PCR protocols will depend on the quality and quantity of the starting chromatin and the efficiency of linker ligation. After amplification, carefully quantitate the amplicon with UV-spectrophotometry and store at -20^{o}C.

9. To prepare more DNA, if needed for labeling, 10 ng of the master stock of the amplicons of the experimental IP, the IgG IP, and the input sample can be amplified for 10 cycles, which should give 1-2 ug of DNA. DNA is purified using the Qiaquick PCR purification kit and then quantitated .

10. Using 10 ng of the experimental IP, the IgG IP, and the input amplicons, a standard PCR reaction is performed using primers corresponding to a known positive control (the same as used to verify that the original ChIP assays had been successful). Do not proceed if the positive control is not enriched in the experimental amplicon relative to the input amplicon or if the IgG amplicon shows non-specific enrichment.

III. CHROMATIN AMPLICON LABELING

For our hybridization experiments we routinely compare the enrichment of the ChIP sample relative to the input chromatin. This provides several advantages: first, this comparison gives a direct (but non-linear) indication as to how robustly a given sequence is enriched during the ChIP procedure. Secondly, by comparing the IP labeled with Cy5 to the input chromatin labeled with Cy3, we generate a ratio that helps reduce bias

introduced by microarray printing errors. Finally, using two color fluorescence reduces the number of arrays required for each experiment.

There are two primary ways to label any given chromatin sample: direct vs. indirect. Direct labeling involves denaturing the chromatin, random priming and polymerizing new duplex DNA with a nucleotide analogue that has the cyanine fluorescent dye conjugated to it. Because these fluorescent dyes are bulky, the polymerase will incorporate the conjugates at different rates than it would a natural nucleotide. This leads to labeling bias; for example, Cy5 tends to incorporate more readily than does Cy3. This has to be taken into account when co-hybridizing both fluorophores onto the same array, and the reciprocal experiment must be done to control for these differences where the test and control samples are each labeled in separate experiments with each dye. One can also directly label the amplicons by including the cyanine-dye conjugated nucleotides directly in the PCR reaction; again because of labeling bias, this method also requires reciprocal experiments in which the samples are labeled with the other dye.

For these reasons, our laboratory primarily utilizes an indirect means of labeling the immunoprecipitated chromatin. Indirect means of labeling incorporates a non-fluorescent nucleotide analogue such as aminoallyl dUTP or biotinylated dUTP, followed by conjugation of the fluorophore to the incorporated nucleotide analogue. Because these small conjugates are incorporated into both the test and reference amplicons equally, this method helps to eliminate the incorporation biases that occur when directly labeling with Cy5-dUTP and Cy3-dUTP. The cyanine dyes can be purchased as an NHS-ester conjugate (Amersham, RPN5661), which will covalently bind to the aminoallyl

nucleotide analogues, or as streptavidin conjugates which forms a tight complex with the biotinylated chromatin. We routinely utilize aminoallyl labeling of our chromatin samples, followed by cyanine dye coupling via the protocol outlined below. In addition the aminoallyl means of amplicon labeling has proved more cost-effective than using one of the direct labeling methods.

Day 1

! 1. 3 X 200 ng of the experimental IP amplicon, 3 X 200 ng of the IgG IP amplicon, and 6 X 200 ng of the input amplicon are each vacuum desiccated to complete dryness with heat (12 total samples) and then each is resuspended in 33 ul molecular biologic grade H_20. Then, 30 ul of 2.5X random primer buffer (Invitrogen, 18094-011) is added to each.

2. The chromatin is denatured by holding at 95°C for 5 min on a PCR block, and then is immediately placed on ice for 3 min. The labeling reaction is initiated by adding 7.5 ul of the 10X dNTP mix (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 0.35 mM dTTP), 1.8 ul of 10 mM aminoallyl-dUTP (Sigma, A0410; make a 0.1 M stock in H_20 and store at –20°C), and 2.5 ul of high concentration Klenow (40U/ul, Invitrogen, 18094-011) to each, and holding at 37°C for two hours.

! 3. The excess nucleotides are removed with the Qiaquick PCR purification kit, eluting with molecular biologic grade H_2O, as the cyanine dyes that are subsequently coupled are sensitive to Tris base. It is very important at this stage to remove as much of the unincorporated aminoallyl-dUTP as possible, because it will bind with the NHS-ester

cyanine dyes and reduce the labeling efficiency of the chromatin. The samples are then dried with heat in a speed-vac.

! 4. To couple the NHS-ester cyanine dye to the chromatin, the aminoallyl-labeled chromatin is resuspended in 4.5 ul H_20, and the cyanine dye is resuspended in 4.5 ul of 0.1 M $NaHCO_3$ (pH 9). The tubes are vortexed and spun several times to ensure that all the samples are dissolved. Next, the chromatin and Cy dye are combined and incubated at room temperature in the dark for 1.5 hours, agitating and microfuging the samples every 15 min of the coupling. We typically couple the Cy5 dye to the experimental IP (or the IgG) amplicon, and Cy3 dye to the input reference amplicon, but these could be reversed as well.

! 5. 35 ul of 100 mM Na-acetate (pH 5.2) are added to lower the pH of the solution to allow the chromatin to bind to the Qiaquick columns, and then H_20 is added to a total of 100 ul for each sample.

6. We have found that removal of the unincorporated dye is quite efficient with Qiaquick PCR columns, and we follow the Qiagen protocol except that we again elute with 50 ul of water instead of the EB buffer provided with the kit.

7. Determination of the minimum amount of labeled DNA required for any given set of CpG microarrays needs to be empirically determined and this depends greatly on the quality of the microarrays that are produced. If one wants a measure of the success of the labeling reaction, the total dye incorporation can determined by measuring the absorbance of the entire sample at 650 nm for Cy5 and 550 nm for Cy3. To determine the number of pmol of incorporated Cy5 or Cy3, the A650 reading for Cy5 and A550 reading for Cy3 are taken on a spectrophotometer after dye conjugation to the aminoallyl

labeled amplicons. The pmol incorporation of dye is calculated by using the following formula: (A650)(ul of solution)(dilution factor)/(0.25 for Cy5, or 0.15 for Cy3)= pmol of Cy dye incorporated. By determining the pmol incorporation necessary to give high quality data on the arrays, one can avoid wasting arrays due to poorly labeled samples. At this stage we also measure the amount of labeled DNA. After aminoallyl labeling of the amplicons, one can ascertain the success of the incorporation of the nucleotide analogue if the measured amount of the DNA is amplified at least 20-30 fold over the starting amount. However, we have found that starting the labeling procedure with 200 ng of labeled chromatin is the least amount of chromatin one can use that will give a good signal/noise ratio on the array. If the arrays are not giving good median intensities per spot (i.e. >500), one can start with two or three times as much amplicon (i.e. 3 X 200ng of Ab, 3 x 200 ng IgG, and 3 X 200ng total input). Because the cyanine dyes are easily photobleached, it is best to perform all steps under minimal light exposure conditions, both during and after the cyanine dyes have been conjugated. The cyanine dyes are also extremely susceptible to degradation by air humidity, so they should be stored dry with a desiccant at all times. Dry the labeled chromatin with heat in a speed-vac and store it dry at –20°C until ready for hybridization. The Cy5 dye is especially susecptable to degradation from various sources including ambient light and ozone.

IV. PREPARATION OF THE ARRAYS

CpG islands, found in or near approximately 50% of human promoters, are identified by three primary characteristics: they are more than 200 bp long, have over 50% GC composition, and retain an observed/expected ratio of CpG dinucleotides greater

than 0.6.[28] The CpG library used on the microarrays was originally prepared in the laboratory of Bird[29] and consists of approximately 60,000 independent clones which represent CpG islands which were not hypermethylated in human male genomic DNA. Approximately 8000 clones representing hypomethylated CpG islands were arrayed by the Huang lab[30] onto slides for high-throughput study of DNA methylation changes in human cancer. We have successfully used these arrays to identify genomic binding sites for E2F6, E2F4, E2F1, and Rb[21, 22]. However, these clones were not sequence verified and therefore cultures of each putative positive target clone had to be grown and DNA prepared and sequenced to determine the location on the genome of the identified CpG island. To circumvent this limiting aspect of the CpG arrays, a new version of the arrays is now in production. For these new arrays, the clones are sequenced prior to spotting. This prior knowledge of the identity of each clone will greatly speed the data analysis and target gene identification steps.

The CGI library[29] was obtained from the United Kingdom Human Genome Mapping Project (http://hgmp.mrc.ac.uk/geneservice/reagnets/) as frozen bacterial cultures (methylation tolerant *E. coli* strain XL1-Blue MRF') habouring individual clones in the pGEM-5Zf(-) vector. The clone set consists of approximately 12,200 independent clones stored in 96-well plates. The entire clone set was amplified in using a 96-well format. The cultures were stabbed and transferred into a total volume of 100 ul containing 4U *Taq* DNA polymerase, 10 mM Tris pH8.4, 50 mM KCl, 1.5 mM $MgCl_2$, 800 uM dNTPs and 0.2 pM each primer. The PCR mixture was held at 94ºC for 2 min and then cycled 40 times at 94ºC for 30 s, 53.5ºC for 30 s, and 72ºC for 1 min, followed by holding at 4ºC after the final cycle. 2.5 ul of each reaction was electrophoresed on a 2

% agarose gel to check the yield and quality of PCR products. The remaining reaction volume was filtered through Unifilter 800 filter plates (Whatman, 7700-2803) and the purified PCR products were quantified by spectrophotometric measurements at 260 and 280 nm. Then approximately 2 ug of DNA from each clone was transferred into 384-well polyfiltronic plates (Whatman, 7701-5101) using Evolution P3 liquid-handling robotics (Perkin Elmer). The PCR products were subsequently dried by overnight centrifugation in vacuum dessicator and resuspended in 10 ul of 3X SSC for array printing. The CpG microarrays were printed from 32 384-well plates containing around 0.2 ug/ul of DNA in 3X SSC in the VersArray ChipWriter Pro System (Bio-Rad, 169-0006) using a 48-pin (SMP-3 pins, ArrayIT, product # SMP3) configuration. Each clone was spotted once on GAPS II coated slides (Corning, 40003). The printed arrays were processed following manufacturer's protocol (Corning) and kept in a dessicator for later use[1].

V. HYBRIDIZATION OF THE ARRAYS

Day 1

! 1. The CpG-island microarray is placed in the hybridization chamber probe side up. A 24X60 mm Lifterslip (Erie Scientific, 25x60I-2-4789) is carefully placed over the area containing the DNA probes to create a uniform space between the array and coverslip.

[1] Contact the Microarray Centre, University Health Network, Toronto, ON M5G2C4 at orders@microarrays.ca for information concerning the distribution of the human CpG arrays.

! 2. The Cy5 and Cy3 labeled chromatin are each resuspended in 15 ul of 1.0 ug/ul CoT-1 DNA (Invitrogen, 15279-011), and the two are then combined. CoT-1 DNA is included to bind to repeat elements in the chromatin which helps to prevent non-specific binding to the array. The tube should be vortexed several times to completely dissolve the chromatin and then spun in a microfuge. 70 ul of the hybridization solution (Genisphere Buffer 6, cat#100V600) is added without vortexing to avoid bubble formation. The mixture is denatured at 95°C for 2 min, and then held at 37°C for 30 min to allow the CoT-1 DNA to hybridize with repeats. The mixture is then applied to the microarray by carefully adding the solution to one end of the array and allowing it to wick underneath the Lifterslip. Note that the size of the Lifterslip and the final volume of the hybridization solution may need to be adjusted to match the printed area of the microarray.

3. We utilize a dual hybridization chamber (Genemachines, HYB-03**)** and a water bath for the hybridization to maintain a constant temperature. It is important that the hybridization chamber not directly contact the bottom of the water bath because it will cause inconsistent heat transfer. We typically use a small piece of foam to separate the chamber from the metal bottom of the water bath. The arrays are hybridized overnight at 60°C for up to 18 hours.

Day 2

11. To remove the coverslip, the microarray(s) are inverted in a glass dish filled with 1X SSC and 0.1% SDS preheated to 50°C. If no drying has occurred overnight, the Lifterslip will immediately slide off.

12. The arrays are then agitated in 1X SSC, 0.1% SDS at 50°C for 5 min, followed by agitation in 1X SSC, 0.1% SDS at room temperature for 5 min. Finally, the arrays are washed with 0.2X SSC for 5 min at room temperature to remove residual SDS (which autofluoresces), and then dried by centrifugation at 600 rpm for 5 min in a 50 ml conical tube.

13. The arrays are scanned immediately using an Axon 4000B scanner (Axon Instruments), or an Agilent scanner, which has the advantage of automatic LOWESS normalization.

References

1. V.R. Iyer, C.E. Horak, C.S. Scafe, D. Botstein, M. Snyder, P.O. Brown. Nature. 409, 533 (2001).

2. J.D Lieb, X. Liu, D. Botstein, P.O. Brown. Nat. Genet. 28, 327 (2001).

3. B. Ren, F. Robert, J.J. Wyrick, O. Aparicio, E.G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T.L. Volkert, C.J. Wilson, S.P. Bell, R.A. Young. Science. 290, 2306 (2000).

4. H.H. Ng, F. Robert, R.A. Young, K. Struhl. Genes Dev. 16, 806 (2002).

5. M. Damelin, I. Simon, T.I. Moy, B. Wilson, S. Komili, P. Tempst, F.P. Roth, R.A. Young, B.R. Cairns, P.A. Silver. Mol. Cell. 9, 563 (2002).

6. J.J. Wyrick, J.G. Aparicio, T. Chen, J.D. Barnett, E.G. Jennings, R.A. Young, S.P. Bell, O.M. Aparicio. Science. 294, 2357 (2001).

7. T.I. Lee, N.J. Rinaldi, F. Robert, D.T. Odom, Z. Bar-Joseph, G.K. Gerber, N.M. Hannett, C.T. Harbison, C.M. Thompson, I. Simon, J. Zeitlinger, E.G. Jennings, H.L.

Murray, D.B. Gordon, B. Ren, J.J. Wyrick, J. Tagne, T.L. Volkert, E. Fraenkel, D.K. Gifford, R.A. Young. Science. 298, 799 (2002).

8. P. Kapranov, S.E. Cawley, J. Drenkow, S. Bekiranov, R.L. Strausberg, S.P. Fodor, T.R. Gingeras. Science. 296, 916 (2002).

9. B. Ren, H. Cam, Y. Takahashi, T. Volkert, J. Terragni, R.A. Young, B.D. Dynlacht. Genes Dev. 16, 245 (2002).

10. A.S. Weinmann, P.S. Yan, M.J. Oberley, T.H. Huang, P.J. Farnham. Genes Dev. 16, 235 (2002).

11. J. Wells, P.S. Yan, M. Cechvala, T.H. Huang, P.J. Farnham. Oncogene. In press. (2002).

12. I.P. Ioshikhes, M.Q. Zhang. Nat. Genet. 26, 61 (2000).

13. S.H. Cross, J.A. Charlton, X. Nan, A.P. Bird. Nat. Genet. 6, 236 (1994).

14. P.S. Yan, C.M. Chen, H. Shi, F. Rahmatpanah, S.H. Wei, T.H. Huang. J. Nutr. 132, 2430S (2002).

15. A.S. Weinmann, P.J. Farnham. Methods. 26, 37 (2002).

16. J. Wells, C.R. Graveel, S.M. Bartley, S.J. Madore, P.J. Farnham. Proc. Natl. Acad. Sci. U S A. 99, 3890 (2002).

17. M.T. Mcmanus, P.A. Sharp. Nature Rev. Gen. 3, 737 (2002).

18. H. Shi, P.S. Yan, C.M. Chen, F. Rahmatpanah, C. Lofton-Day, C.W. Caldwell, T.H. Huang. Cancer Res. 62, 3214 (2002).