High-resolution nucleosome mapping of targeted regions using BAC-based enrichment

Erbay Yigit[1], Quanwei Zhang[2], Liqun Xi[2], Dan Grilley[1], Jonathan Widom[1,†],
Ji-Ping Wang[2,*], Anjana Rao[3,4] and Matthew E. Pipkin[3,4,*]

[1]Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208, USA, [2]Department of Statistics, Northwestern University, Evanston, IL 60208, USA, [3]Division of Signaling and Gene Expression, The La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037, USA and [4]Program in Cellular and Molecular Medicine, Children's Hospital and Harvard Medical School, Boston, MA 02115, USA

ABSTRACT

We report a target enrichment method to map nucleosomes of large genomes at unprecedented coverage and resolution by deeply sequencing locus-specific mononucleosomal DNA enriched via hybridization with bacterial artificial chromosomes. We achieved ~10 000-fold enrichment of specific loci, which enabled sequencing nucleosomes at up to ~500-fold higher coverage than has been reported in a mammalian genome. We demonstrate the advantages of generating high-sequencing coverage for mapping the center of discrete nucleosomes, and we show the use of the method by mapping nucleosomes during T cell differentiation using nuclei from effector T-cells differentiated from clonal, isogenic, naïve, primary murine CD4 and CD8 T lymphocytes. The analysis reveals that discrete nucleosomes exhibit cell type-specific occupancy and positioning depending on differentiation status and transcription. This method is widely applicable to mapping many features of chromatin and discerning its landscape in large genomes at unprecedented resolution.

INTRODUCTION

Nucleosomes are the fundamental repeating subunits of chromatin, and they consist of an octamer of histone proteins (H2A, H2B, H3 and H4) around which ~147 bp of DNA is tightly wrapped. Adjacent nucleosomes are spaced by a short ~20–50-bp segment of 'linker' DNA (1,2). Binding sites for sequence-specific DNA-binding factors are less accessible when the DNA containing the sites is wrapped in nucleosomes. In turn, this regulated accessibility influences the ability of transcription factors to recruit and mobilize the enzymatic machinery of replication, repair, recombination and transcription.

The rules that govern nucleosome organization *in vivo* are under intense scrutiny (3–6). Deciphering these rules depends on methods to measure nucleosome occupancy across the genome, to assign their center locations at high-resolution *in vivo* and to quantify how they are remodeled during cell activation and differentiation. The most widely used approach involves treating native chromatin with micrococcal nuclease (MNase), isolating DNA protected from digestion by the histone octamer and subjecting the protected DNA to high-throughput sequencing. Computational methods are used to calculate nucleosome maps based on an occupancy curve resulting from alignment of the sequenced reads to a genome (7,8). Much of our understanding of nucleosome positioning *in vivo* derives from model organisms, such as yeast, in which adequate sequencing coverage to map nucleosomes genome-wide at high-resolution can be achieved economically because they have relatively small genomes (~250 times smaller than humans and mice). But these model organisms typically lack the complex structure encoded in mammalian genomes. Most nucleosomes have not been mapped in mammalian genomes at high resolution because the cost to achieve sufficient whole-genome sequencing coverage of nucleosomal DNA in such large genomes remains prohibitively expensive, and only a small fraction of nucleosomes seem to be strongly positioned (6). As a result, much of what is known from whole-genome sequencing of human nucleosomal DNA derives

from the minority of nucleosomes that are strongly positioned and patterns of nucleosome occupancy that result after summarizing all nucleosome-derived reads from regions surrounding aligned structural features (e.g. transcriptional start sites or DNase I hypersensitive sites) in bins of genes, grouped based on their mRNA expression (6,9). Thus, nucleosome organization of mammalian genomes has mainly been defined by the average ensemble of nucleosome occupancy from many loci, rather than high-resolution knowledge of discrete nucleosomes within individual genes.

We envisioned target enrichment of mononucleosomal DNA as a strategy to achieve high enough sequencing coverage of nucleosome DNA to map discrete nucleosomes and to focus on genomic regions that are likely to contain nucleosomes that undergo remodeling during cell differentiation. In this report, we developed a method to enrich mononucleosomal DNA from specific loci using hybridization to bacterial artificial chromosomes (BACs), by modifying a previously reported direct genomic selection method (10), introducing a multiplexing step and adapting the method to massive parallel sequencing technologies. We call this approach BEM-seq (bacterial artificial chromosomes enriched mono-nucleosomal DNA sequencing).

MATERIALS AND METHODS

BAC DNA isolation and labeling

Appropriate BAC clones were identified via the UCSC genome browser and were purchased from Children's Hospital of Oakland Research Institute. BAC DNA was isolated from 1 l of culture, purified by chromatography (Qiagen) and analyzed for identity and quantity by restriction enzyme mapping using pulse-field gel electrophoresis (PFGE) and spectrophotometry as described previously (11). All BAC preparations used for enrichment were confirmed to be $\geq$80% intact BAC DNA by PFGE. For each hybridization reaction, BAC DNA was labeled with a commercial nick translation kit (Roche cat. 10976 776 001). The reaction (total 20 µl) was set-up in a polymerase chain reaction (PCR) tube on ice as follows: 120 ng of BAC DNA, 5 µl of Biotin-dUTP/dNTP mixture, 2 µl of 10× nick translation buffer, 0.5 µl of α-^{32}P [dCTP] (6000 Ci/mmol) and 2 µl of enzyme mixture. Biotin-dUTP/dNTP mixture: Biotin-16-dUTP (0.8 mM) 0.6 volume, dTTP (0.8 mM) 2.4 volume, dATP (0.8 mM) 3.0 volume, dCTP (0.8 mM) 3.0 volume and dGTP (0.8 mM) 3.0 volume. Nick translation was carried out at 15°C for 1 h in a thermal cycler. The reaction was stopped by addition of ethylenediaminetetraacetic acid (EDTA). The labeled BAC DNA was purified by size-exclusion chromatography (Bio-Rad, Micro Bio-Spin P-30). Biotinylated-dUTP (Enzo Life Sciences) incorporation was confirmed indirectly based on radioactivity: About 10% of the labeled BAC DNA was incubated with magnetic streptavidin beads in streptavidin-binding buffer [10 mM Tris–HCl (pH7.5), 1 mM EDTA and 1 M NaCl] at room temperature, and labeled DNA was collected on a magnet. The relative amount of radioactivity in biotinylated DNA

captured with the magnetic beads versus that left in the supernatant was measured using a Geiger counter. If at least four times greater radioactivity was found in the beads relative to the supernatant, the BAC DNA was considered to be successfully biotinylated.

Mice, cells and T-cell differentiation

Naïve CD4 T and CD8 T cells were isolated from 6- to 16-week-old OT-II and P14 T-cell receptor (TCR) transgenic (Tg) mice, respectively, that were maintained on the $Tcra^{-/-}$ background to eliminate T cells co-expressing endogenously rearranged TCRs. All mice were maintained in specific pathogen-free facilities and used according to protocols approved by the Immune Disease Institute and Harvard Medical School and La Jolla Institute for Allergy and Immunology animal care and use committees. T cells were isolated from single cell suspensions prepared from pooled spleen and lymph nodes, by positive (CD4 cells) or negative (CD8 cells) selection (Dynal). For T-cell differentiation, 1×10^7 naïve OT-II CD4 T cells, or naïve P14 CD8 T cells, were activated for 48 h in 10 ml complete T-cell media supplemented with hamster αCD3 (2C11) and αCD28 (37.51) each at 1 µg/ml, in T25 flasks pre-coated with goat anti-hamster capture antibody. For T helper type 1 (Th1) cultures, media was supplemented with interleukin (IL)-12 (10 ng/ml) and αIL-4 (11B11; 10 µg/ml); for T helper type 2 (Th2) cultures, media was supplemented with IL-4 (supernatant, ~1000 U/ml), αIFNγ (XMG1.2; 5 µg/ml) and αIL-12 (3 µg/ml). After 48 h, Th1 and Th2 cultures were resuspended, counted and diluted to 5×10^5 cells/ml in fresh Th1 or Th2 media supplemented with 20 U/ml recombinant human IL-2 (rhIL-2; NIH bioresource). Th1 and Th2 cultures were expanded every 24 h by counting and readjusting cells to 5×10^5 cells/ml, using Th1 or Th2 media with rhIL-2 until day 4, and thereafter using media containing only rhIL-2. After the initial 48 h of stimulation, CD8 T cells were removed from the TCR stimulation and diluted to 5×10^5 cells/ml in fresh media containing 100 U/ml of rhIL-2 and were counted and readjusted to 5×10^5 cells/ml 100 U/ml of rhIL-2 media every 24 h (12). The differentiation status of T-cell cultures was monitored by flow cytometry using labeled antibodies recognizing CD4 (RM4-5), CD8a (53-6.7), CD25 (PC61) CD44 (IM7), CD62L (Mel-14), IL-7Rα (A7R34), T-bet (4B10) and Eomes (Dan11Mag) (Figure 1A and data not shown). Cytokine gene transcription was elicited pharmacologically by stimulating Th1 and Th2 cells at a density of 1×10^6/ml in the presence of 10 nM phorbol mystirate acetate (PMA) and 1 µM ionomycin for 6 h; brefeldin A was included for the final 2 h of stimulation. Re-stimulated cells were harvested, washed in phosphate-buffered saline (PBS), fixed in 2% paraformaldehyde, washed again in PBS and then permeabilized with 0.25% saponin in PBS supplemented with 1.0% Fetal Bovine Serum (FBS) (saponin-staining buffer). Permeabilzed cells were stained in saponin-staining buffer using labeled αIL-4 (11B11) and αIFNγ (XMG1.2) antibodies. Stained cells were washed in saponin-staining buffer before finally re-suspending PBS/ 1.0% FBS for analysis (data not shown).

Nuclei isolation and MNase digestion

Nuclei were prepared as previously described (11,13), with the following modifications detailed later in the text, although other methods to digest native chromatin that require less handling time are likely to work as well are known (14), and they might offer some advantages for mapping nucleosomes accurately. Exponentially growing lymphocyte cultures were harvested in a chilled centrifuge (4°C), washed, concentrated and adjusted to 2×10^7 cells/ml in ice cold Ca^{2+}/Mg^{2+} free PBS. One volume of cells was lysed on ice for 10 min by adding four volumes of lysis buffer [12.5 mM Tris (pH 7.4), 45 mM KCl, 6.25 mM $MgCl_2$, 375 mM sucrose and 0.125% nonidet P-40, supplemented with 2 mM benzaminidine and one complete EDTA-free protease inhibitor cocktail/50 ml]. Nuclei were pelleted at 600*g* for 7 min in a pre-cooled swinging bucket rotor at 4°C. Pelleted nuclei were resuspended in nuclei wash buffer [10 mM Tris (pH 7.4), 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$ and 300 mM sucrose] using one-tenth volume relative to the initial volume at lysis. Resuspended pellets were pooled 2:1 to concentrate, and the final wash volume was adjusted to half the initial lysis volume with nuclei wash buffer. The nuclei were pelleted again as above, and supernatants were completely aspirated. For MNase (Sigma) digestion, pelleted nuclei were resuspended in 1 ml MNase buffer [10 mM Tris (pH 7.4), 15 mM NaCl, 60 mM KCl, 150 µM spermine and 500 µM spermidine] and then were adjusted to 2×10^7 nuclei/ml in MNase buffer, on ice. MNase digestion was performed at 37°C in the presence of 1 mM $CaCl_2$, and the incubation time was titrated to identify optimal digestion for each nuclei preparation. Incubation times resulting in chromatin digested into ~80% 147-bp fragments (core nucleosome) that were clearly resolved from 167-bp fragments (chromatosome) with limited evidence of intranucleosomal cleavage, were selected for a large preparative digestion. The digestion was stopped with a final concentration of 10 mM EDTA and 1% sodium dodecyl sulfate (SDS), RNase A (20 µg/ml) was added and the lysates were incubated at 37°C for 30 min. The final concentration of NaCl was increased to 0.5 M, Proteinase K was added to a final concentration of 50 µg/ml and the lysates were incubated at 55°C for at least 1 h. DNA was extracted with phenol/chloroform, precipitated with ethanol and resuspended in TE buffer.

Gel extraction and purification of mononucleosomal DNA

Digested chromatin DNA was resolved on 3.3% agarose gels (NusieveGTG, Lonza) in 1× TBE buffer until 147- and 167-bp DNA bands were clearly resolved. The 147-bp DNA band was carefully excised from the agarose gel, crushed and soaked in a buffer containing 300 mM sodium acetate, 1 mM EDTA and 0.05% SDS for 24–48 h at room temperature with gentle agitation on a rocker. Agarose was removed from the crush and soak buffer by Amicon Ultrafree-CL filter (Millipore), and filtrates were concentrated with Amicon-15 Ultra filter devices (30 kDa cut-off; Millipore) and were purified by QIAquick Spin Columns using buffer PB. Commercial gel extraction kits containing strong chaotropic reagents that melt AT rich sequences easily were found to substantially denature a large fraction of mononucleosomal DNA (data not shown), and thus were not used to reduce potential bias towards nucleosomal DNA with high GC content (15).

Preparation of SOLiD library

About 6 µg DNA was treated with a commercial DNA end-repair enzyme cocktail (Lucigene Cat No 40037-2) to make DNA ends ligatable. Repaired DNA ends were ligated to SOLiD P1 and P2 adaptors with the Quick Ligase Kit (New England Biolabs cat. no M2200L). For some of the samples, repaired DNA ends were ligated to custom barcoded SOLiD P1 adaptors (see Table 2) for multiplexing. Ligation products were separated on 6% native polyacrylamide gels, excised, extracted using the crush and soak method and purified. The gaps on unligated strands were repaired by DNA polymerase I (New England Biolabs cat. no M0209L) in the presence of dNTPs at 16°C for 30 min. After each aforementioned step, the DNA was purified using QIAquick PCR Purification Kit (Qiagen). DNA concentration in the libraries was determined by absorbance at 260 nm in TE buffer using a Nanodrop (ThermoFisher Scientific).

Enrichment of targets by BAC hybridization

Labeled BAC DNA was transferred to a 200-µl PCR tube and lyophilized using a speedvac without applying any heat. To suppress repeats during hybridization, DNA pellets were resuspended in 5 µl of TE buffer containing 10 µg of mouse *cot-1* DNA (Invitrogen) and overlaid with mineral oil. Labeled BAC DNA was denatured at 96°C for 5 min and incubated at 65°C for an additional 15 min in a thermal cycler, and then 5 µl of 2× hybridization buffer [1.5 M NaCl, 40 mM sodium phosphate buffer (pH 7.2), 10 mM EDTA (pH 8), 10× Denhardt's and 0.2% SDS] was added to the DNA, and the mixture was incubated for 6 h at 65°C. In a separate 200-µl PCR tube, 5 µl (1–2 µg) of nucleosome library DNA was denatured at 96°C for 5 min, and then incubated for an additional 15 min at 65°C in a thermal cycler, and then 5 µl of 2× hybridization buffer was added to the denatured library DNA. The entire volume of denatured nucleosome library DNA was added to the *cot-1* suppressed labeled BAC DNA, and the mixture was incubated at 65°C for ~72–80 h. Fifty microliters of streptavidin-coated magnetic beads (Invitrogen) was washed with binding buffer and resuspended in 150 µl streptavidin-binding buffer in a low-retention microcentrifuge tube (Eppendorf cat. no 022 431 021). The hybridization mix was directly added to the pre-washed magnetic beads, and binding was carried out on a rotator at room temperature for 30 min. The magnetic beads were washed once at 25°C for 15 min in 1× SSC with 0.1% SDS and three times at 65°C for 15 min in 0.1× SSC with 0.1% SDS. All the washes were carried out in 1 ml of wash buffer on a heat block in a microcentrifuge tube with occasional mixing. After washes were completed, the hybridized nucleosomal DNA was denatured from the BAC hybrids by adding 100 mM

Table 1. Quantification of BAC-based target enrichment of mononucleosomal DNA using colony sequencing

Cell type	BAC	Size	Number of on-target clones[a]	Total number of clones	On-target clones (%)	Fold-enrichment
CTL	Prf1	224 kb	67	95	71	8600
Th1	Prf1	224 kb	24	30	80	9700
Th2	Prf1	224 kb	25	30	83	10 000
[b]Seven different cell types	Nine different BAC clones	1.87 Mb	27	45	60	810 [c]

Fold-enrichment = (number of on-target clones/number of total sequenced clones) × [the size of mouse Genome (bp)/the size of BAC clone (bp)].
[a]On-target clones = clones mapping to target BAC used for enrichment.
[b]Nucleosome DNA from seven different cell types were barcoded during library preparation, and each library was enriched by hybridization with a mixture of nine different BAC clones. The enriched DNA was cloned and sequenced using standard methods. Of 45 sequenced clones, 27 matched regions in the nine BACs used for selection.
[c]The average size of one BAC within the nine used for multiplexed selection was 207.5 kb, corresponding to an average of 7300-fold enrichment per BAC clone.

NaOH at room temperature for 10 min, and the magnetic beads were removed from the elution solution at room temperature. The elution solution was transferred to a new microcentrifuge tube, neutralized with 100 µl of 1 M Tris (pH 7.5) buffer and desalted by a size-exclusion column (Bio-Rad P-30 column). DNA was amplified by PCR (15–18 cycle) using P1 and P2 PCR primers, gel extracted and submitted for SOLiD 4 sequencing. To quantify target enrichment, we blunt-end cloned an aliquot of the BAC-enriched PCR products from each enrichment (Invitrogen ZeroBlunt PCR cloning kit) and sequenced the inserts from antibiotic-resistant colonies using standard methods (Table 1). Enrichment was calculated as follows: Fold-enrichment = [genome size (bp)/BAC size (bp)] × (number of sequenced clones aligned to target BAC)/(total number of sequenced clones).

Calculation of center-weighted nucleosome occupancy score

For the single-end data, the average reads length is unknown. To construct the nucleosome reads occupancy, we need to know where the projected nucleosome center is based on each tag. However, the average reads length in different BACs or different regions on the same BAC may differ. Without adjustment of the average reads length, the correlation analysis (described later in the text) of the nucleosome occupancy can be biased. Suppose position i is the nucleosome center, then the reads on the Watson strand originating from this nucleosome may start from $i\text{-}k$, and the reads from Crick strand may start from $i+k$ for some integer k. For practical purpose, we only consider Watson tags that reside 63–83-nt upstream and Crick tags 63–83-nt downstream, or equivalently, the complete reads length could vary between 127 and 167 bp if both ends of the DNA insert had been sequenced. Let w_i and c_i be the observed tag frequency at position i on Watson and Crick strands, respectively. For any single-end tag on Watson position i, as its paired-end on Crick is unknown, the tag frequency in the region from position $i+126$ to $i+166$ on the Crick strand, i.e. $c_{i+126}, \ldots, c_{i+166}$, can be used to estimate the probability of that the paired-end start at the corresponding positions. Thus, if we observe w_i tags at position i on the Watson strand, then the projected centers of the

nucleosomes attributed to the w_i tags are positions $i+63, i+63.5, \ldots, i+83$ with corresponding reads frequency weight $w_i c_{i+k}/\sum_{j=i+126}^{i+166} c_j$ for $k = 126, \ldots, 166$.. The weight at half positions will be split into halves and added to neighboring integer positions. Likewise, for c_i tags at position i on the Crick strand, the projected nucleosome center positions are $i - 63, i - 63.5, \ldots, i - 83$ with weights $c_i w_{i-k}/\sum_{j=i-126}^{i-166} w_j$ for $k = 126, \ldots, 166$. Combing both strands, the expected reads frequency that is centered at position i (if both ends had been sequenced) can be calculated as

$$
\begin{aligned}
R_i = & \sum_{k=63}^{83} w_{i-k}\left(c_{i+k}/\sum_{j=i-k+126}^{i-k+166} c_j\right) \\
& +0.5\sum_{k=63}^{82} w_{i-k}\left(c_{i+k+1}/\sum_{j=i-k+126}^{i-k+166} c_j\right) \\
& +0.5\sum_{k=63}^{82} w_{i-k-1}\left(c_{i+k}/\sum_{j=i-k+125}^{i-k+165} c_j\right) \\
& +\sum_{k=63}^{83} c_{i+k}\left(w_{i-k}/\sum_{j=i+k-126}^{i+k-166} w_j\right) \\
& +0.5\sum_{k=63}^{82} c_{i+k}\left(w_{i-k-1}/\sum_{j=i+k-126}^{i+k-166} w_j\right) \\
& +0.5\sum_{k=63}^{82} c_{i+k+1}\left(w_{i-k}/\sum_{j=i+k-125}^{i+k-165} w_j\right)
\end{aligned}
\tag{1}
$$

We define the center-weighted nucleosome occupancy score at position i as follows:

$$
o_i = \sum_{|j-i|\leq 60} R_j G(i-j; 20),
\tag{2}
$$

where G stands for a Gaussian density with SD = 20. The Gaussian weight function covers ± 60 bp of the projected nucleosome center and decays towards two ends. The advantage of using this weighting function is to avoid spikes in the reads occupancy curve frequently observed in the linker region because of overlap of reads arising from two neighboring nucleosomes when a uniform weight function is applied.

For the paired-end data, the number of reads that are centered at each position is directly observed. Therefore, Equation (2) can be readily applied to calculate the center-weighted nucleosome occupancy.

Nucleosome center DNA alignment

Based on the center-weighted occupancy scores of the paired-end data, we selected the strongest peaks sequentially based on peak height and peak steepness, requiring that two selected neighboring peaks be spaced by at least 120 bp. The selected peak positions were treated as the nucleosome centers to examine the nucleosome-specific dinucleotide signal (e.g. the AA/TT/TA/AT frequency). The selected peaks represent the approximate center positions of the nucleosomes, but the precision could be undermined by the well-known sequence preference of MNase. In addition, we searched for a sequencing read of length 147 bp nearest to the peak position in the ±5-bp region because reads with length of ~147 bp tend to be better representatives of the true nucleosome positions compared with reads that are much shorter or longer than 147 bp. If no such sequence existed, we further searched for sequences of lengths 148, 146, 149, 145 and 150 bp sequentially within ±5-bp region of the peak until the first sequence was identified. The center of the identified sequence was treated as the nucleosome center to generate the AA/TT/TA/AT frequency plot. If no such sequence was identified in the ±5-bp region, the peak position itself was treated as the true nucleosome center to generate the alignment.

Correlation analysis

Let $O_k = \{o_{1k}, o_{2k}, ..., o_{nk}\}$, $k = 1, ..., K$ be the center-weighted nucleosome occupancy score series in a genomic region of length n for condition 1 through K as calculated previously. We calculated the pairwise Spearman rank-based correlation between any two conditions for a given genomic region.

RESULTS

Generation of mononucleosome DNA libraries for NG sequencing and BAC enrichment

To map mammalian nucleosomes, we used a series of well-established culture systems that drive the differentiation of naïve T cells into defined effector T-cells subsets, a process that is associated with pronounced chromatin and transcriptional changes (11,16). Thus, we purified clonal populations of naïve CD4 and CD8 T cells (Figure 1A) from inbred TCR Tg mice and differentiated these precursor cells into Th1 and Th2 cells, and Cytotoxic T lymphocyte (CTL), respectively (Figure 1B). Compared with purified T cells from human donors that have been used for nucleosome mapping (6,9), which are polyclonal, genetically disparate and notoriously heterogeneous functionally and phenotypically (17), purified mouse T-cells differentiated in culture are likely to have the least biological variability in nucleosome organization between individual cells (and alleles) within the population because

they are clonal, isogenic and strongly polarized towards a specific phenotype. Th1, Th2 and CTL cells each differentially remodel chromatin structure in the *Ifng*, *Il4* and *Prf1* loci, and this underlies the specific transcription of each gene in the different cell types (Figure 1B). We isolated nuclei from each of the three cell types and incubated them with MNase for increasing amounts of time to exhaustively cleave linker DNA and produce fragments protected by the histone octamer; DNA was extracted from digested chromatin and was carefully resolved on 3.3% low-melting point 1× TBE agarose gels (Figure 1C). Conditions that resulted predominantly in canonical mononuclesomal DNA fragments (i.e. 147 bp) with limited evidence of intranucleosomal cleavage were selected for library preparation (Figure 1C). Mononucleosome DNA fragments were excised from the gel, purified, end-repaired and ligated to adaptors compatible with Applied Biosystems SOLiD sequencing to generate mononucleosome DNA libraries (Figure 1D). In some cases, we used P1 adapters customized with 6-bp barcodes to sequence different samples simultaneously as a mixture (see 'Materials and Methods' section). Notably, both careful attention to the extent of MNase digestion and the subsequent size selection step reduced recovery of smaller DNA fragments derived from protection by other non-histone DNA-binding proteins, such as transcription factors and chromatin remodeling factors, or particularly unstable nucleosomes, which could otherwise complicate subsequent computation of nucleosome maps (18). To enrich nucleosomal DNA from targeted loci, we hybridized the linkered libraries with individual biotinylated BAC clones covering the *Il4*, *Ifng* and *Prf1* loci in separate reactions (Figure 1D). Mononucleosomal DNA annealed to the BACs was captured using streptavidin beads; non-specific DNA was washed away; and the captured nucleosomal DNA was eluted from the biotinylated BACs by alkali denaturation, then purified and subjected to 15–18 cycles of PCR with adaptor-specific SOLiD P1- and P2-specific primers (Figure 1D). We quantified enrichments by cloning the enriched mononucleosome libraries, sequencing inserts from 30 to 100 clones using standard methods and aligning the sequences with the mouse genome (Figure 1E). Based on 10 independent enrichment experiments using BAC clones spanning multiple different loci, we found that an average of 65% of clones mapped to the corresponding BAC regions used for selective capture, which equated to >8600-fold enrichment for a 224 kb BAC clone after a single hybridization (Table 1 and data not shown).

Next, we multiplexed the BAC enrichment to capture sequences from multiple loci in a single hybridization. Mononucleosome DNA libraries prepared from each cell type were hybridized with mixtures of up to nine different BAC clones to simultaneously capture nucleosomes covering ~1.9 Mb of the genome. We sequenced the resulting libraries with standard methods and confirmed that each target locus was specifically enriched after hybridization with multiplexed BACs. The total enrichment of all regions covered by the nine different BACs in the multiplex was 810-fold, and enrichment of the specific
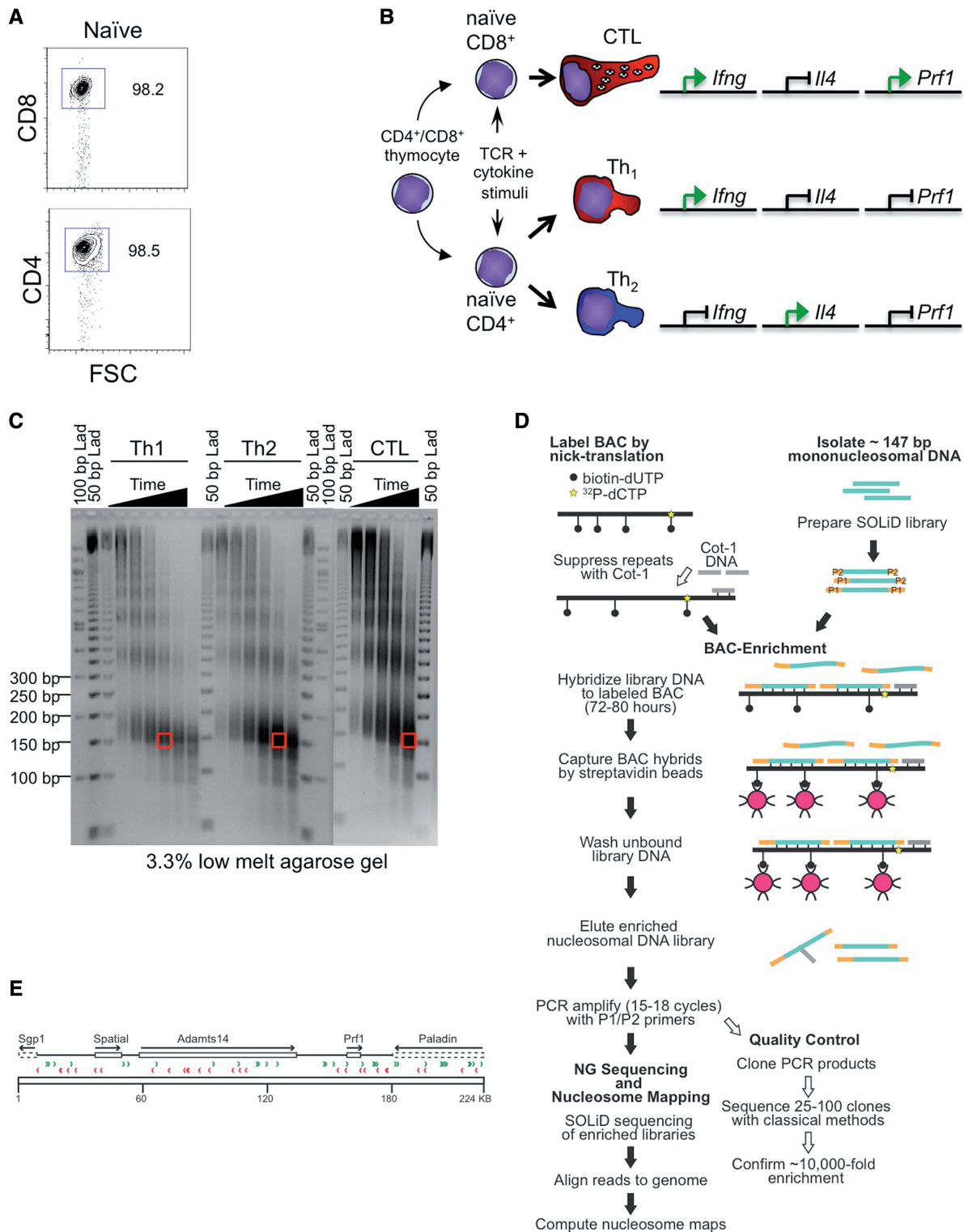
Figure 1. Homogeneous populations of T cells were used for nuclei preparations, isolation of mononuclesomal DNA and BEM-seq. (**A**) Flow cytometry was used to confirm the purity of naïve CD4 and CD8 T cells after isolation. The percentage of CD4+ and CD8+ T cells of all events is shown. (**B**) Schematic shows the developmental relationship of CD4 and CD8 T cells, and it depicts their differentiation into classical Th cell and CTL subsets and their characteristic pattern of transcription of the *Ifng*, *Il4* and *Prf1* genes (green arrows indicate transcriptionally active genes). (**C**) Nuclei isolated from each different cell type were incubated with MNase for increasing time (wedges above blots) before the DNA was extracted and resolved on 3.3% low-melt agarose gels in 1× TBE. Mononucleosomal DNA (~147 bp) was excised from the gel, and the DNA was extracted to generate mononucleosomal DNA libraries. Note that only DNA from time points in which there was limited evidence of intranucleosome cleavage were selected for library preparation. (**D**) Schematic of BEM-seq method shows the work-flow. (**E**) Alignment of Sanger sequenced clones to the *Prf1* BAC used for enrichment. Standard dye-terminator methods were used to sequence inserts from 95 colonies after blunt-end cloning DNA from the enriched and amplified nucleosome library. Seventy percent of the clones matched perfectly to the target region (67 of 95). A representative alignment is shown for a nucleosome library that was enriched using clone RP24-323A22, which covers the murine *Prf1* gene. Colored arrows indicate the strands to which the sequences aligned.

Table 2. Quantification of BAC-based target enrichment of mononucleosomal DNA using SOLiD paired-end sequencing

Cell type	Barcode sequence[a]	Number of unique tags	Number of non-unique tags	Number of unique on-target tags	On-target tags (%)	Number of BAC clones used for enrichment[b]
CTL	None	5 191 686	471 033	4 180 213	81	7[(3)]
CTL BR	CGCTAG	4 003 477	466 044	2 053 070	51	9
Th1	None	5 257 325	420 894	4 233 849	81	7[(4)]
Th1 BR	TGTGGG	4 227 774	448 603	2 970 319	70	9
Th2	None	1 197 925	85 950	829 754	69	7[(3)]
Th2 BR	GAATGT	3 782 956	391 195	2 938 927	78	9

[a]Nucleosome DNA from biological replicas (BR) was ligated to barcoded P1 and original P2 adaptors. We added 6 bp of unique sequence (barcode) to the 3′-end of the SOLiD P1 adaptor as follows: CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGATNNNNNN (N being barcode sequence). Barcoding enabled sequencing mixtures of multiple libraries from different conditions simultaneously on one sector of an octet slide in a single flow cell, to significantly reduce sequencing costs.
[b]Superscript numbers indicate that several independent enrichments with different BACs were performed before the enriched libraries were combined and subjected to parallel sequencing. For example, 7[(3)] indicates that 3 independent hybridizations that collectively included a total of seven different BACs were used for enrichment, and then the enriched DNA from each hybridization was pooled for sequencing.

region covered by each individual BAC was on average >7000-fold, similar to when a single BAC was used during an independent enrichment (Table 1). These results were based on nuclei preparations from eight different cell types and two biological replicates (Table 2 and data not shown). Together, these data show that nucleosomal DNA from distinct loci covering megabase regions can be enriched simultaneously, efficiently and reproducibly.

BEM-seq maps genuine nucleosomes with very high precision

To generate nucleosome maps, we sequenced each enriched mononucleosome DNA library independently using both single-end and paired-end SOLiD sequencing chemistries. Alignment of single-end (35-bp tags) or paired-end (35- and 25-bp tags) reads to the murine genome (mm9) confirmed that enrichment was consistent with results from standard sequencing methods (Table 2). The aligned reads were used to calculate the center-weighted nucleosome occupancy score at each position of enriched regions (see 'Materials and Methods' section for details). Reads from paired-end sequencing were mapped more efficiently than single-end reads, and more importantly, paired-end sequencing resulted in better mapping accuracy of the nucleosome centers than single-end sequencing because both ends of the protected DNA were known (data not shown and see 'Materials and Methods' section). Using the paired-end data, we examined whether called peaks were likely to be the center of nucleosomes by 'center aligning' the ±73-bp region at the identified occupancy peak position and plotting the resulting frequency of AA/AT/TT/TA dinucleotides (see 'Materials and Methods' section for details). A clear 10-bp dinucleotide periodicity was evident in the alignment (Figure 2A), a hallmark of sequences associated with nucleosomes *in vivo* because they tolerate sharp bending of DNA around the histone octamer (1,19–21). These data show that targeted enrichment of mononucleosomal DNA by hybridization to BAC clones can be used to map individual nucleosomes from specific megabase genomic regions of the mammalian genome.

We achieved up to ~1000-fold coverage (i.e. uniquely mapped reads × 147 bp/total bp in BACs used for selection) after multiplexed enrichment and paired-end sequencing in some experiments. The very high coverage afforded by BEM-seq enabled us to examine how sequencing coverage could affect mapping precision using a simulation study. Approximately 300 000 uniquely aligned paired-end reads of defined length (137–157 bp) (Supplementary Figure S1) were used to map nucleosomes in 210 kb of the *Il4* locus. With the full data set (~210-fold coverage), 791 peaks were identified on the occupancy curve, and the peak positions were treated as putative 'true' nucleosome centers. A fixed fraction of reads from the full data set were randomly sampled, and the putative nucleosomes were identified again in the same way as for the original sample. We plotted the frequency of distances between the nucleosome positions in the original map and the simulated maps from different sampling fractions (Figure 2B). At ~20-fold coverage, only about half of the peaks mapped back to within ±4 bp of the original nucleosome centers, and the distance from the original centers rapidly increased for the remaining 50% of peaks. In contrast, at 100-fold coverage, >80% of peaks could be mapped to within ±4 bp of the original nucleosome centers. These data show that increased sequencing coverage achieved with BEM-seq improves the precision of mapping individual nucleosome centers.

BEM-seq discerns differential nucleosome organization in distinct T-cell lineages

To examine whether nucleosomes mapped by BEM-seq could be used to discern biologically relevant changes in nucleosomes, we inspected nucleosome organization of the transcribed region of *Prf1* in CTL, Th1 and Th2 cells. *Prf1* is strongly transcribed in CTL, whereas it is weakly transcribed, or silent, in Th1 and Th2 cells (11). Surprisingly, the Spearman correlations of the overall nucleosome occupancy scores throughout the transcribed region of *Prf1* in any pairwise comparison of the three different cell types were very high. However, the correlation between Th1 and Th2 cells was higher than the
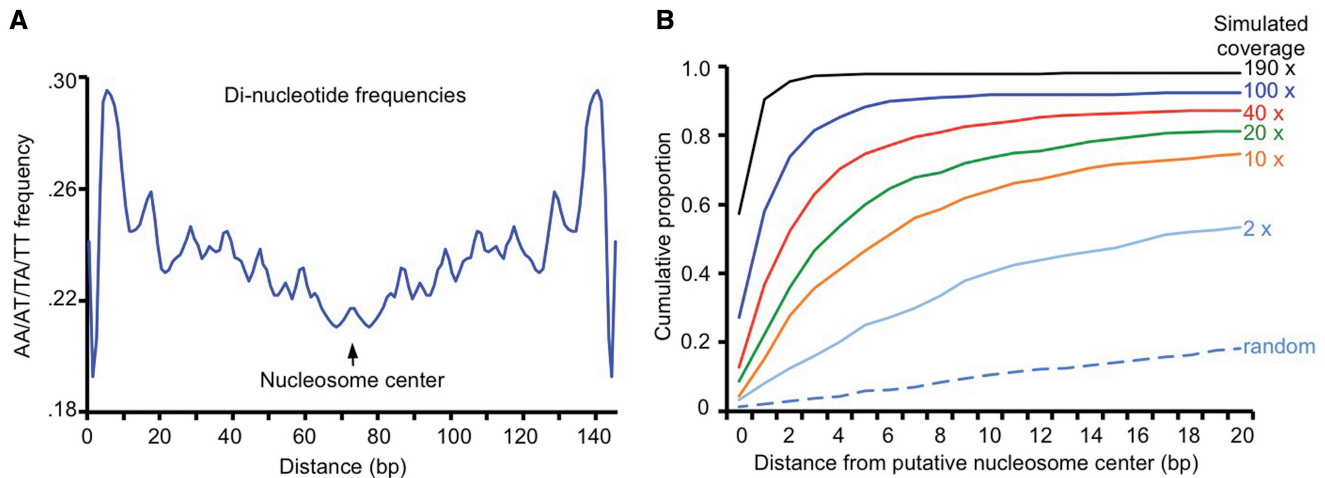
Figure 2. BEM-seq maps genuine nucleosomes and high sequencing coverage improves mapping precision. (**A**) The frequency of dinucleotides in sequences of center-aligned peaks from the *Il4, Ifng* and *Prf1* loci in Th1 cells shows a characteristic 10-bp periodicity typical of sequences bound to nucleosomes *in vivo*. (**B**) The effect of sequencing coverage on mapping nucleosome centers precisely was determined by comparing nucleosome center locations mapped at very high coverage, relative to their locations in maps calculated from the same data in which reduced sequencing coverage was simulated. The cumulative proportion of distances between the 'true' putative centers of 791 nucleosomes in maps calculated with the entire data set, and nucleosome centers in maps calculated with progressively fewer reads is shown. The approximate fold-coverage for each simulated map (colored lines) is indicated. Note that as simulated sequencing coverage decreases, more nucleosomes were mapped with larger deviations relative to their 'true' locations.

correlations between CTL and either Th1 or Th2 cells, indicating that nucleosomes were distinct in CTL relative to either Th cell lineage. This result was true whether nucleosome maps were calculated based on single-end sequencing or paired-end sequencing in different biological replicates (Table 3). Although overall most nucleosomes seemed to be organized similarly in all three cell types, discrete nucleosomes exhibited strikingly different occupancy or positioning in CTL relative to both Th1 and Th2 cells. These nucleosomes were primarily located at the transcriptional start site (TSS), upstream of the promoter and in the second intron of *Prf1* (Figure 3A), and were located near previously identified DNase I hypersensitive (DHS) sites in CTL and Th1 cells (11). Specifically, several nucleosomes in the promoter and −1-kb region showed lower nucleosome occupancy in CTL relative to Th1 and Th2 cells. In addition, a strongly positioned '+1' nucleosome was located immediately downstream of the TSS in CTL, but the occupancy of this nucleosome was strikingly lower in both Th1 and Th2 cells (Figure 3A, lower panel). In addition, the center location of the +2 nucleosome in Th1 and Th2 cells was shifted towards the TSS compared with its location in CTL. Thus, differential organization of specific nucleosomes was correlated with strong transcription of *Prf1* in CTL and weak or inactive transcription in Th cells.

We also examined the transcribed aspects of the *Ifng* and *Il4* genes. Analogous to observations in the *Prf1* gene, there was overall similarity of the nucleosome pattern in both the *Il4* and *Ifng* genes in CTL, Th1 and Th2 cells, whereas the occupancy and the positioning of certain discrete nucleosomes were clearly different between the three cell-types. These differentially re-modeled nucleosomes were located near to the TSS

Table 3. Spearman correlation of nucleosomes across *Prf1* in Th1, Th2 and CTL

Chemistry	Condition	Th1	Th2	CTL	Th1 BR	Th2 BR	CTL BR
Single-end	Th1	1.0	**0.9297**	**0.6799**	0.7640	0.7515	0.6432
	Th2		1.0	**0.6870**	0.7182	0.7718	0.6185
	CTL			1.0	0.7686	0.7480	0.8516
Paired-end	Th1 BR				1.0	**0.9336**	**0.8918**
	Th2 BR					1.0	**0.8567**
	CTL BR						1.0

Nucleosomes were mapped in two biological replicates (denoted BR) using single-end sequencing for the first replicate and paired-end sequencing for the second replicate. Note that in both cases, correlations between Th1 and Th2 cells were higher than between CTL and either Th1 or Th2 cells (bold text).

region and overlapped known DHS sites (22,23). In the *Ifng* gene, the +1 nucleosome exhibited high occupancy in all three cell types, whereas its center location was specifically repositioned in Th1 and Th2 cells relative to CTL (Figure 3B, upper panel). Interestingly, the repositioned +1 nucleosome did not seem to be associated with alterations in any of the immediately adjacent downstream nucleosomes (Figure 3B, lower panel). In contrast, at the *Il4* locus in Th2 cells, a delocalized nucleosome signal covered the region where a +1 nucleosome would presumably be located downstream of the TSS, and the immediate downstream nucleosomes were substantially different between Th1 and Th2 cells (Supplementary Figure S2). Taken together, these data demonstrate that BEM-seq resolves differential changes in individual nucleosomes near the transcribed aspects of differentially transcribed genes during effector T-cell differentiation.
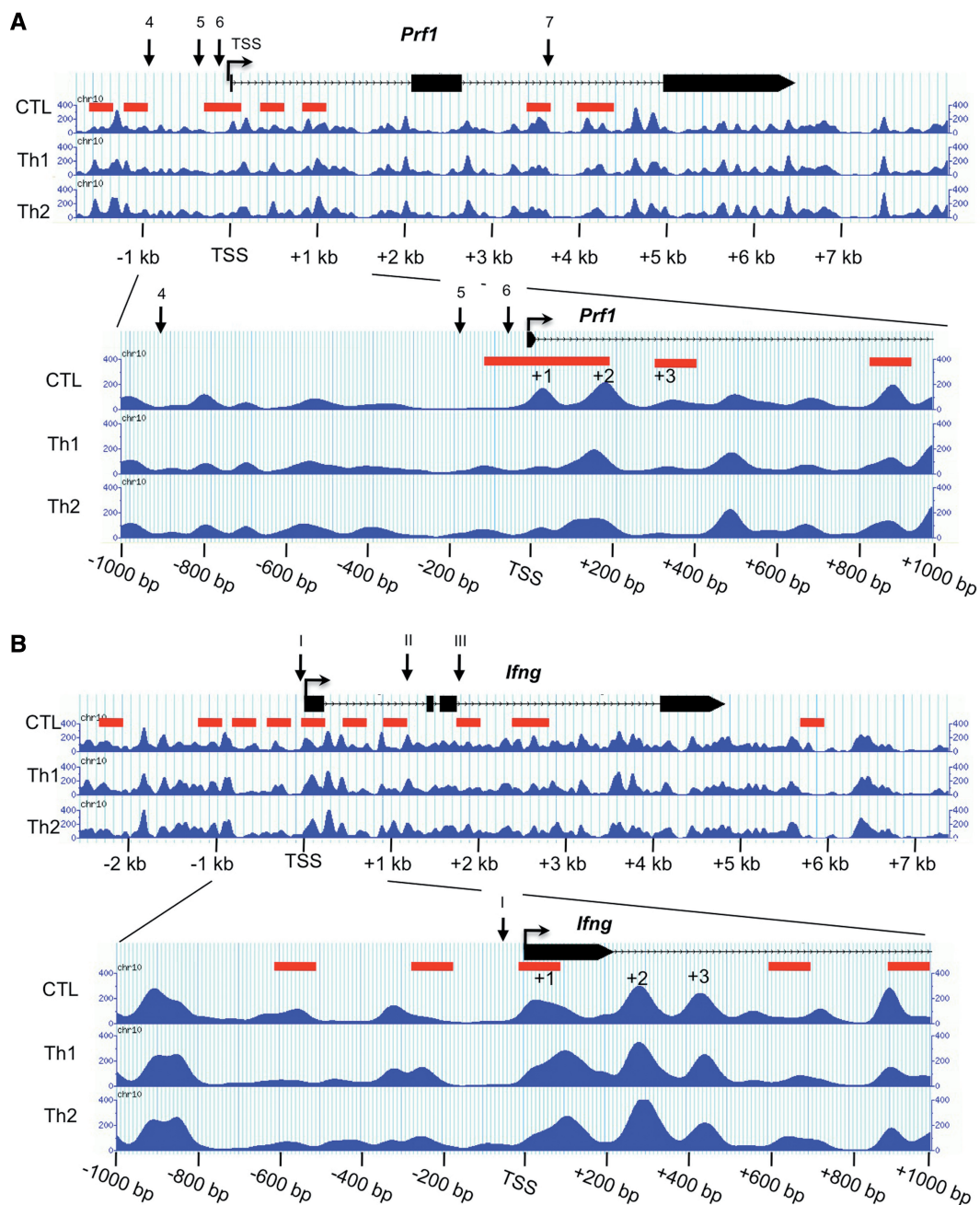
Figure 3. Nucleosome organization around TSSs is cell type specific. (**A**) BEM-seq identifies changes in specific nucleosomes in *Prf1*. Nucleosome tracks across 10 kb of *Prf1* (chr10: 60758820–60768819, mm9) in CTL, Th1 and Th2 cells and the position of DNase I hypersensitive sites (numbered arrows) are shown. Nucleosomes in the +1, +2 and +3 positions are denoted. Note, depletion of nucleosomes at the TSS and −1-kb region, a strongly positioned +1 nucleosome immediately downstream of the TSS and remodeled nucleosome positions near DHS 7 in CTL relative to Th1 and Th2 cells. The red bars highlight nucleosomes that are altered between at least two cell types. (**B**) Nucleosome tracks and the position of DNase I hypersensitive sites (numbered arrows) across 10 kb of *Ifng* (chr10: 117875525–117885524) in CTL, Th1 and Th2 cells are shown. Note the downstream repositioning of the +1 nucleosome in Th1 and Th2 cells relative to CTL, but a lack of attendant remodeling of nucleosomes in the +2 and +3 positions.

DISCUSSION

To map nucleosomes at high-resolution in large genomes and focus on relevant loci, we developed an approach called BEM-seq, which is based on enriching mono-nucleosome DNA derived from specific regions by hybridization with BACs, and subjecting the enriched DNA to parallel sequencing. We achieved enrichments of up to

10 000-fold for a given locus (~200 kb) after a single hybridization, and we required <18 cycles of PCR to amplify the libraries post-enrichment. To reach comparable enrichment, a previously reported method for direct genomic selection with BACs required two sequential rounds of selection and up to 60 cycles of PCR (10). Thus, the BEM-seq approach enriches targets more efficiently; therefore, enriched libraries are likely to be

influenced less by bias resulting from extensive PCR compared with previous methods. Solution hybridization with synthetic oligonucleotides has also been used for target-enrichment and sequencing, and it has noted skewing of coverage based on nucleotide content (24). The skewing might have been because of inefficient hybridization of particular sequences, but it might also be ascribed to artifacts in capture sequences arising during oligonucleotide synthesis, combined with known biases inherent to preparation and sequencing of next-generation libraries (15). In any case, the BEM-seq approach is likely to be subject to similar biases during hybridization and library preparation, but result in a faithful or better representation of original sequences because of the fidelity of BAC-derived sequences relative to synthesized DNA. Synthetic oligonucleotides have been used to enrich megabase regions of the genome by hybridization, and we also demonstrated that BEM-seq enriches many loci simultaneously, by 'multiplexing' several different BACs in the same hybridization. This enabled high-resolution mapping of nucleosomes across ~1.9 Mb in eight different cell types in a single experiment. Therefore, the BEM-seq method is scalable similar to using synthetic oligonucleotides, yet is much less expensive and could conceivably be used to interrogate most or all genes that are differentially expressed in cells exposed to different activation stimuli or during a particular differentiation process. Thus, BEM-seq is a reliable and economical method for enriching mononuclesome DNA from targeted, yet expansive regions of genomes to sequence the associated nucleosomes at very high coverage.

The increased nucleosome sequencing coverage afforded by BEM-seq improved the precision of mapping nucleosomes based on the reads data, and this is likely to have translated into knowing their positions *in vivo* more accurately. The apparent positions of nucleosomes mapped by MNase digestion may differ from their actual positions *in vivo*, especially when sequencing coverage is low, because MNase prefers to cleave particular sequences (25), and there is dynamic behavior of nucleosomes (18) leading to random variation in read lengths and center positions of DNA protected from MNase cleavage. We found that paired-end sequencing and 100-fold coverage mapped a much larger proportion of nucleosomes precisely than when using ~20-fold coverage (which is roughly comparable with current whole-genome single-end sequencing of mononucleosome DNA in humans). We found that at ~20-fold coverage, the location of a large fraction of nucleosome peaks deviated substantially from their likely 'true' centers. We presume those nucleosomes whose center locations did not deviate drastically at low coverage likely comprised the minority of nucleosomes that were 'well positioned', consistent with findings from whole-genome sequencing of mononucleosome DNA from human T cells (6). Beyond mapping nucleosomes that are 'well positioned', our data indicate that at least 100-fold coverage of mammalian genomes with paired-end sequencing may be necessary to map most nucleosomes precisely. Finally, given that recent studies indicate MNase-derived bias minimally affects discerning the actual positions of nucleosomes

in vivo (21,26), the increased precision afforded by BEM-seq likely improves mapping nucleosomes with respect to their real positions *in vivo*.

Nucleosome organization assayed by whole-genome sequencing of mononucleosome DNA is typically visualized in aggregate plots as patterns of nucleosome occupancy surrounding aligned genomic landmarks, such as TSSs. The average pattern of nucleosomes at the TSS of active genes consists of a nucleosome-depleted region upstream of the TSS, a strongly positioned '+1' nucleosome immediately downstream of the TSS and phasing of sequential downstream nucleosomes; in genes that are not expressed highly, this pattern is less amplified (6,9). Our high-resolution mapping of nucleosomes in the *Prf1*, *Ifng* and *Il4* genes with BEM-seq indicated that nucleosome organization around TSSs is gene specific and much more complex than as viewed in the average pattern of many genes. We noted several important distinctions. First, both transcriptionally active and inactive genes exhibited a nucleosome-depleted region upstream of the TSS, suggesting that this configuration is not programmed by transcriptional competence itself. Second, although transcriptional changes correlated with changes in organization of the '+1' nucleosome, in certain cases, occupancy of the +1 nucleosome was altered (see the *Prf1* TSS), whereas in others, the location of +1 nucleosome was changed, or it was not strongly positioned in the first place (see the *Il4* and *Ifng* TSSs). Interestingly, differential positioning of the +1 nucleosome did not correspond to strong changes in the positioning of adjacent nucleosomes as would be predicted if placement of the +1 nucleosome enforced statistical positioning and downstream phasing of nucleosomes (5). Thus, high-resolution mapping of discrete nucleosomes reveals finite architecture that is not discernable in the ensemble average. Our data suggest that BEM-seq is likely to be a powerful approach to examine how DNA sequence, transcription factor occupancy and chromatin remodeling enzymes precisely place individual nucleosomes in mammalians.

BEM-seq is particularly suited for comparative mapping of the same genomic regions under different conditions or in different cell types. Analysis of the *Prf1* locus revealed previously unappreciated details of chromatin accessibility in *cis*-acting domains. We previously found that the DNase I hypersensitivity pattern of the *Prf1* promoter and immediate 5′ flank in CTL and Th1 cells did not explain the 20-fold higher expression of *Prf1* in CTL compared with Th cells; both cell-types exhibited the same DNase I hypersensitivity pattern (11). Here, using BEM-seq, however, we detected obvious and reproducible cell type-specific differences in nucleosome organization of these regions in CTL relative to Th cells. Notably, depletion of nucleosomes in CTL relative to Th cells correlated with high-RNA Polymerase II occupancy at the *Prf1* TSS, and high occupancy of Stat5 and eomesodermin transcription factors at the known Stat5 dependent enhancer at −1 kb, suggesting that these factors may be responsible for the altered nucleosome pattern (12,27). We found analogous changes in the *Ifng* and *Il4* loci. Therefore, results based on mapping nucleosomes with BEM-seq provide distinct and specific regulatory information

compared with other measures of chromatin accessibility, such as DNase I hypersensitivity.

In summary, BEM-seq is likely a broadly applicable approach for probing many aspects of the chromatin landscape within relevant regions of large and complex genomes. In addition to mapping nucleosomes, combining BAC-based enrichment with next-generation sequencing is likely to be ideal for quantifying the occupancy and location of other DNA bound factors, the apposition of distal *cis*-elements and the location of non-coding transcripts that may be unrepresented or invisible in the ensemble average of whole-genome measurements. Thus, applying the BEM-seq approach together with ChIP-seq and chromatin conformation assays, and RNA-seq is likely useful for resolving, in unprecedented detail, the interactions that underlie chromatin dynamics and transcription.

SUPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1 and 2.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Segal,E. and Widom,J. (2009) What controls nucleosome positions? *Trends Genet*, **25**, 335–343.
2. Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251–260.
3. Kaplan,N., Moore,I.K., Fondufe-Mittendorf,Y., Gossett,A.J., Tillo,D., Field,Y., LeProust,E.M., Hughes,T.R., Lieb,J.D., Widom,J. *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, **458**, 362–366.
4. Zhang,Y., Moqtaderi,Z., Rattner,B.P., Euskirchen,G., Snyder,M., Kadonaga,J.T., Liu,X.S. and Struhl,K. (2009) Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat. Struct. Mol. Biol.*, **16**, 847–852.
5. Zhang,Z., Wippo,C.J., Wal,M., Ward,E., Korber,P. and Pugh,B.F. (2011) A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science*, **332**, 977–980.
6. Valouev,A., Johnson,S.M., Boyd,S.D., Smith,C.L., Fire,A.Z. and Sidow,A. (2011) Determinants of nucleosome organization in primary human cells. *Nature*, **474**, 516–520.
7. Valouev,A., Ichikawa,J., Tonthat,T., Stuart,J., Ranade,S., Peckham,H., Zeng,K., Malek,J.A., Costa,G., McKernan,K. *et al.* (2008) A high-resolution, nucleosome position map of C elegans reveals a lack of universal sequence-dictated positioning. *Genome Res.*, **18**, 1051–1063.
8. Zhang,Z. and Pugh,B.F. (2011) High-resolution genome-wide mapping of the primary structure of chromatin. *Cell*, **144**, 175–186.
9. Schones,D.E., Cui,K., Cuddapah,S., Roh,T.Y., Barski,A., Wang,Z., Wei,G. and Zhao,K. (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell*, **132**, 887–898.
10. Bashiardes,S., Veile,R., Helms,C., Mardis,E.R., Bowcock,A.M. and Lovett,M. (2005) Direct genomic selection. *Nat. Methods*, **2**, 63–69.
11. Pipkin,M.E., Ljutic,B., Cruz-Guilloty,F., Nouzova,M., Rao,A., Zuniga-Pflucker,J.C. and Lichtenheld,M.G. (2007) Chromosome transfer activates and delineates a locus control region for perforin. *Immunity*, **26**, 29–41.
12. Pipkin,M.E., Sacks,J.A., Cruz-Guilloty,F., Lichtenheld,M.G., Bevan,M.J. and Rao,A. (2010) Interleukin-2 and inflammation induce distinct transcriptional programs that promote the differentiation of effector cytolytic T cells. *Immunity*, **32**, 79–90.
13. Pipkin,M.E. and Lichtenheld,M.G. (2006) A reliable method to display authentic DNase I hypersensitive sites at long-ranges in single-copy genes from large genomes. *Nucleic Acids Res.*, **34**, e34.
14. Stewart,A.F., Reik,A. and Schutz,G. (1991) A simpler and better method to cleave chromatin with DNase 1 for hypersensitive site analyses. *Nucleic Acids Res.*, **19**, 3157.
15. Quail,M.A., Kozarewa,I., Smith,F., Scally,A., Stephens,P.J., Durbin,R., Swerdlow,H. and Turner,D.J. (2008) A large genome center's improvements to the Illumina sequencing system. *Nat. Methods*, **5**, 1005–1010.
16. Ansel,K.M., Djuretic,I., Tanasa,B. and Rao,A. (2006) Regulation of Th2 differentiation and Il4 locus accessibility. *Annu. Rev. Immunol.*, **24**, 607–656.
17. Catalina,M.D., Sullivan,J.L., Brody,R.M. and Luzuriaga,K. (2002) Phenotypic and functional heterogeneity of EBV epitope-specific CD8+ T cells. *J. Immunol.*, **168**, 4184–4191.
18. Henikoff,J.G., Belsky,J.A., Krassovsky,K., MacAlpine,D.M. and Henikoff,S. (2011) Epigenome characterization at single base-pair resolution. *Proc. Natl Acad. Sci. USA*, **108**, 18318–18323.
19. Satchwell,S.C., Drew,H.R. and Travers,A.A. (1986) Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.*, **191**, 659–675.
20. Wang,J.P. and Widom,J. (2005) Improved alignment of nucleosome DNA sequences using a mixture model. *Nucleic Acids Res.*, **33**, 6743–6755.
21. Brogaard,K., Xi,L., Wang,J.P. and Widom,J. (2012) A map of nucleosome positions in yeast at base-pair resolution. *Nature*, **486**, 496–501.
22. Agarwal,S. and Rao,A. (1998) Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. *Immunity*, **9**, 765–775.
23. Balasubramani,A., Shibata,Y., Crawford,G.E., Baldwin,A.S., Hatton,R.D. and Weaver,C.T. (2010) Modular utilization of distal cis-regulatory elements controls Ifng gene expression in T cells activated by distinct stimuli. *Immunity*, **33**, 35–47.
24. Tewhey,R., Nakano,M., Wang,X., Pabon-Pena,C., Novak,B., Giuffre,A., Lin,E., Happe,S., Roberts,D.N., LeProust,E.M. *et al.* (2009) Enrichment of sequencing targets from the human genome by solution hybridization. *Genome Biol.*, **10**, R116.
25. Dingwall,C., Lomonossoff,G.P. and Laskey,R.A. (1981) High sequence specificity of micrococcal nuclease. *Nucleic Acids Res.*, **9**, 2659–2673.
26. Allan,J., Fraser,R.M., Owen-Hughes,T. and Keszenman-Pereyra,D. (2012) Micrococcal nuclease does not substantially bias nucleosome mapping. *J. Mol. Biol.*, **417**, 152–164.
27. Zhang,J., Scordi,I., Smyth,M.J. and Lichtenheld,M.G. (1999) Interleukin 2 receptor signaling regulates the perforin gene through signal transducer and activator of transcription (Stat)5 activation of two enhancers. *J. Exp. Med.*, **190**, 1297–1308.