Published in final edited form as:

Anal Chem. 2010 March 15; 82(6): 2480–2487. doi:10.1021/ac9028642.

Single Molecule Epigenetic Analysis in a Nanofluidic Channel

Benjamin R. Cipriany[§], Ruqian Zhao[†], Patrick J. Murphy[‡], Stephen L. Levy[□], Christine P. Tan^Δ, Harold G. Craighead^{□,*}, and Paul D. Soloway^{∞,*}

Department of Electrical and Computer Engineering, Department of Nutritional Science, Department of Molecular Biology and Genetics, School of Applied and Engineering Physics, and Department of Biomedical Engineering, Cornell University, Ithaca, New York 14853

Abstract

Epigenetic states are governed by DNA methylation and a host of modifications to histones bound with DNA. These states are essential for proper developmentally regulated gene expression and are perturbed in many diseases. There is great interest in identifying epigenetic mark placement genome-wide and understanding how these marks vary among cell types, with changes in environment or according to health and disease status. Current epigenomic analyses employ bisulfite sequencing and chromatin immunoprecipitation, but query only one type of epigenetic mark at a time, DNA methylation or histone modifications, and often require substantial input material. To overcome these limitations, we established a method using nanofluidics and multi-color fluorescence microscopy to detect DNA and histones in individual chromatin fragments at about 10 Mbp/min. We demonstrated its utility for epigenetic analysis by identifying DNA methylation on individual molecules. This technique will provide the unprecedented opportunity for genome-wide, simultaneous analysis of multiple epigenetic states on single molecules using femtogram quantities of material.

Keywords

Single-molecule; chromatin; epigenetics; epigenomics; DNA methylation; nanofluidics; laser-induced fluorescence; methyl binding domain protein; green fluorescent protein; HeLa cell

Chromatin within eukaryotic cells includes DNA and histone proteins assembled on DNA into the nucleosome¹. The DNA sequence carries the genetic code and controls inheritance of traits, however, reversible covalent modifications to specific DNA sequences and their associated histones can influence how the underlying DNA is utilized and can therefore also control traits^{2,3}. These have been referred to as epigenetic modifications. The most common epigenetic modification to DNA in mammals is methylation and hydroxymethylation of DNA, either of which may be placed on the fifth carbon of the cytosine pyrimidine ring. A host of modifications including methylation, acetylation, ribosylation, phosphorylation, sumoylation, ubiquitylation

*To whom correspondence should be addressed: Harold Craighead hgc1@cornell.edu; (telephone) 607-255-8707; (fax) 607-255-7658 or Paul Soloway soloway@cornell.edu; (telephone) 607-254-6444; (fax) 607-254-6444..

[§]Department of Electrical and Computer Engineering, Cornell University, USA

[†]Key Laboratory of Animal Physiology and Biochemistry, Nanjing Agricultural University, China (Current Address)

[‡]Department of Molecular Biology and Genetics, Cornell University, USA

[□]School of Applied and Engineering Physics, Cornell University, USA

^ΔDepartment of Biomedical Engineering, Cornell University, USA

[∞]Department of Nutritional Science, Cornell University, USA

SUPPORTING INFORMATION AVAILABLE

Additional supporting information, as noted in the text, is available containing detailed experimental protocols, a collection of five figures, and one table. This material is available free of charge via the Internet at <http://pubs.acs.org>.

and citrullination occur at more than 30 amino acid residues of the four core histones within the nucleosome.

Specific changes in epigenetic state that occur genome wide appear to regulate cellular differentiation during development⁴. Perturbations in normal epigenetic state in mature tissues contribute to initiation and progression of cancer and other diseases⁵. Additionally, evidence indicates epigenetic states are influenced by environmental variables including diet⁶, environmental toxins⁷ and maternal behaviors⁸. Given the fundamental role that epigenetic mechanisms play in normal development, environmental responses and how their perturbation affect disease state, there is increasing effort devoted to characterize the human epigenome⁹.

Current methods for epigenomic analysis depend on the epigenetic mark queried. Histone modifications are most commonly detected using chromatin immunoprecipitation (ChIP), in which modification-specific antibodies are used to immunoprecipitate the associated DNA, which is then detected by hybridization to microarray¹⁰ (ChIP-chip) or deep sequencing¹¹ (ChIP-seq). DNA methylation can also be detected by immunoprecipitation using a methylcytosine antibody¹², or with bisulfite sequencing, which offers more comprehensive analysis of DNA methylation states¹³. Genome wide epigenomic analyses using antibodies often use on the order of 10^6 to 10^7 cells. ChIP has been used with few as 100 cells, however, with this few cells the analysis was locus specific and not genome wide¹⁴. A far more significant limitation is encountered when studies seek to determine whether or not two or more epigenetic marks are coincident within the genome. Analysis of each epigenetic mark requires an independent immunoprecipitation. When precipitating chromatin from an ensemble of cells with different antibodies, it is difficult to distinguish true coincidence of the detected marks from the existence of multiple populations within the ensemble, each with a different epigenomic profile. This can be somewhat overcome with sequential ChIP, where the material precipitated by one antibody is re-ChIPed with a second antibody¹⁵. However, these techniques are not amenable to genome wide analysis or for studies in which more than two epigenetic marks are investigated.

Nanofluidic platforms that combine high throughput detection and analysis of single chromatin fragments can overcome the limitations of existing epigenomic methods. Such platforms have been used to size femtogram quantities of DNA¹⁶, to study microRNA gene expression¹⁷, and to quantify mRNA transcript levels¹⁸ and PCR amplification products¹⁹. However, they have not been adapted for use with chromatin, nor used to examine epigenetic marks. Here we established the conditions for flowing and detecting single native chromatin molecules through such a device by analyzing time-coincident fluorescent signatures of both the DNA and histone proteins within the chromatin at high throughput. In addition, we extended our analysis to study an epigenetic mark in naked DNA using conditions that paralleled our chromatin studies and a fluorescently labeled probe that can bind to methylated DNA. Our approach, which we refer to as SCAN (single-chromatin analysis at the nanoscale) is the first demonstration of single-molecule high-throughput epigenetic analysis.

EXPERIMENTAL SECTION

HeLa Cell Chromatin

HeLa cells constitutively expressing green fluorescent protein on histone H2B (H2B-GFP) were provided by Geoffrey M. Wahl at The Salk Institute for Biological Studies, USA. Cell cultures were cultured as monolayers in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% fetal calf serum. Native chromatin fragments were prepared from HeLa cells as described²⁰. For full details see the Supporting Information.

Methyl Binding Domain (MBD1) Protein Synthesis and Labeling

Plasmid for bacterial expression of 1xMBD (pET-1xMBD) was provided by Adrian P. Bird at The Wellcome Trust Centre for Cell Biology, University of Edinburgh, UK. Recombinant His6-tagged MBD1 was purified from 100 mL induced BL21(DE3) cultures on Ni-NTA agarose (Qiagen) using denaturation and on column renaturation cycles in accordance with the manufacturer's instructions, with some modifications. For full details see the Supporting Information.

Lambda DNA Preparation and In-Vitro Methylation

Lambda DNA from phage grown in a methylation deficient host (Promega D1521) was digested with *HindIII* and methylated in vitro with *SssI* methyltransferase, which can methylate all 3,113 CpGs in the 48.5kbp genome. Efficacy of the methylation reaction was assessed by resistance to digestion by the methylation sensitive enzyme *HpaII*.

MBD-DNA Affinity Reaction

In a mixture of methylated and unmethylated DNA suspended at 50 ng/ μ L in 1x Tris Buffered Saline (1x TBS, 50mM TRIS-HCl and 138mM NaCl and 2.7mM KCl, pH 8.0), we performed DNA staining using TOTO-3. Approximately 1 μ g of labeled DNA was then diluted into 20 μ L of buffer containing 1x TBS with 2% bovine serum albumin and 0.1% TritonX-100 (v/v). A volume of 1 μ L MBD1, stored at 280 ng/ μ L, and labeled with AlexaFluor488 was then added to the DNA to perform the binding reaction under conditions of molar excess. The binding reaction occurred for 2 hours at room temperature.

DNA Labeling

We use the cell-impermeant, intercalator TOTO-3 (Invitrogen). The labeling reaction was conducted by mixing the diluted DNA or chromatin and the diluted dye according to the method described by the manufacturer. All samples were prepared with a 1:5 dye to base pair ratio, unless otherwise noted. Following the labeling reaction, samples were protected from light and stored overnight at 4 °C. TOTO-3 exhibits significant fluorescence enhancement upon binding, which alleviated the need for purification to remove unbound dye following the labeling reaction.

Fabrication of Nanofluidic Channels

Nanofluidic channels were fabricated in a fused silica substrate. Projection photolithography (GCA Autostep 200) was used to pattern fluidic channels with a 500 nm critical dimension. This method allowed rapid patterning of 27 fluidic channel arrays, totaling 432 fluidic channels, on a 100 mm diameter wafer. Patterns formed in the developed photoresist were transferred approximately 250 nm into the silica using reactive ion etching (Oxford 80, Oxford Instruments). The fluidic channels were protected with photoresist during subsequent through-wafer drilling using a focused jet of alumina abrasive to form the access ports at the ends of the channels. The wafer surface was cleaned with a hot Piranha solution (3 H₂SO₄ : 1 H₂O₂) and RCA standard clean (5 H₂O : 1 NH₄OH : 1 H₂O₂, heated to 70 °C). Direct touch bonding with a 170 μ m thick coverslip wafer was performed to cap the fluidic channel. A subsequent high-temperature anneal to 1050 °C permanently bonded the stack of fused silica wafers together. An optical-grade epoxy (Norland Products) was used to attach fluid reservoirs to the wafer surface.

Electrokinetic Flow in Nanofluidics

Samples were kept in their respective storage buffer until DNA labeling and/or MBD binding reactions. Following these reactions, samples were serially diluted in 1x Tris-Borate-EDTA

(1x TBE, 89mM TRIS borate and 2mM EDTA, pH 8.0), with additives 0.5% polyvinylpyrrolidone (PVP) measured (w/v) and 0.1% TritonX-100 measured (v/v) (both from Sigma Aldrich, St. Louis, MO). The final dilution suspended the samples at an estimated concentration of 600 pM, nominally 1-2 ng/ μ L for chromatin samples, and about 0.25 ng/ μ L for methylated DNA samples. The polymer additives in the buffer served to limit electroosmotic flow and prevent non-specific interactions with the fluidic channel walls without denaturing proteins. We loaded 50 μ L of the sample solution into the input reservoir of a fluidic channel array and then connected to the negative electrode. The output reservoir contained only the buffer solution and was connected to the positive electrode. Samples were flowed at an applied bias of 50 V for all chromatin samples and 100 V for all methylated DNA samples. Stable electrokinetic flow was established during a pre-flow time of 20 min to ensure steady-state flow conditions had been achieved prior to data collection. Each sample was examined for a total of 15 min, always using the same fluidic channel within the array. Following single molecule detection, the fluidic channel array was rinsed iteratively for a total of 30 min and then checked to verify the absence of fluorescently-labeled molecules, prior to loading the next sample. Fluid channel arrays used with chromatin experiments were not reused with DNA methylation experiments.

Laser Induced Fluorescence Confocal Microscopy

Single molecule fluorescence was observed using an inverted microscope (IX-71, Olympus) equipped with a side laser port. Laser illumination of 330 μ W at 488 nm (Sapphire, Coherent) and 1300 μ W at 635 nm (Cube, Coherent) were overlapped in free-space, incident on a dual-band dichroic mirror (488/647rpc, Chroma Technology) and then focused onto the nanofluidic channel using a 60x, 1.2 numerical aperture water-immersion objective (UPlanSAPO, Olympus), and aided with an electron multiplied CCD camera (Cascade 512B, Photometrics). A dual-band laser notch filter (NF01-488/647, Semrock) attenuated stray laser light and passed single molecule fluorescence. Confocal spatial filtering occurred using a 50 μ m diameter pinhole (901PH, Newport). Two color fluorescence was then chromatically split using a second dichroic mirror (FF560-Di01, Semrock) and then filtered by bandpass fluorescence filters (525/50M and 680/40M, Chroma). Each color of fluorescence was then collected using a 100 μ m diameter core multimode optical fiber (OZ Optics). Photons were detected by avalanche photodiodes (APD) in single photon counting mode (SPCM, Perkin-Elmer) and recorded at 100 kHz using a high-speed correlator (correlator.com) and a personal computer.

Statistical Analysis

Propagated error analysis was performed to evaluate the proportion of bound molecules, chromatin or MBD-DNA, with respect to total DNA. First, we define a window region that encompasses the full width of the Gaussian distribution in a time coincidence histogram. Adjacent to the window region are the sidebands, which were used to characterize the background of uncorrelated molecules. The background contribution in the window region was calculated based upon the uncorrelated molecules per 50 μ s bin in the sidebands, reported as the mean and standard deviation, and then scaled by the number of bins within the window region. The total molecules counted within the window region was summed and reported with a Poisson counting error. The number of bound molecules was then calculated by subtracting the total molecules from the background molecules within the window region and propagating the error. Second, the number of total DNA molecules observed was counted and reported with a Poisson counting error. The ratio of bound molecules and total DNA molecules was then calculated and the errors of each were propagated. As applicable, we plot the average value and error bars that represent the propagated error.

Fitting error analysis was performed to measure the concentration of molecules measured within a nanofluidic channel. The inter-event time separation was plotted as a histogram and

then fitted to single-term exponential decay using Matlab's built-in fitting routine. The fitted mean inter-event time was calculated with a 95% confidence interval. The mean and confidence interval were then evaluated using a Poisson model, to describe molecule occupancy within the inspection volume, to report concentration values.

RESULTS AND DISCUSSION

A Nanofluidic Platform for Single Molecule Analysis

Our strategy for high throughput analysis of single molecules of methylated DNA and chromatin entailed time-resolved detection and spectral identification of fluorescent labels bound to individual molecules containing epigenetic marks of interest. We flowed and detected DNA and chromatin in a solution confined within a nanofluidic channel. These channels reduce the optical excitation volume for fluorescent analysis, thus enabling us to interrogate individual molecules in solutions of relatively high concentration²¹. The nanofluidic channels were fabricated in a fused silica substrate using photolithography and reactive ion etching. Figure 1a shows a single nanofluidic channel made by this process with cross-sectional dimensions of approximately 250 nm wide by 500 nm deep. We formed 27 separate fluidic channel arrays, each with 16 parallel channels, on a 100 mm diameter wafer²². Each array also had access ports with a reservoir at the ends of the channels, which we used to load the samples and to insert electrodes for controlling electrokinetic flow.

To analyze individual molecules of DNA and chromatin flowing through channels, we mounted the silica wafer on a laser confocal microscope and illuminated the samples flowing through an individual channel with two overlapping Gaussian shaped laser profiles, each with a diameter of approximately 1.3 μm (Figure 1b). The laser profiles were larger than the channel width, so that every fluorescent molecule was interrogated with the same illumination profile. The illuminated inspection volume within the channel was 0.16 fL. A burst of fluorescence, associated with the different labeled components, was emitted as each molecule passed through the illumination profile. The colors composing a molecule's signature were separated with optical filters and then detected with avalanche photodiode detectors. A record of the red and green fluorescence bursts were accumulated at 100 kHz onto a personal computer and then analyzed with a custom Matlab algorithm to identify single molecule detection (SMD) events^{18, 23}.

Nucleosome Detection on Native Chromatin

In order for this platform to be a viable tool for epigenomic analysis, we needed to demonstrate that chromatin could be directed through the nanofluidic channel, remain intact, meaning nucleosomes remained on the DNA, and that we could detect fluorescent signatures of the histone and DNA components. For this test we used chromatin extracted from HeLa cells bearing a transgene expressing an H2B - GFP (green fluorescent protein) fusion protein²⁴. H2B-GFP incorporated into nucleosomes allowed the chromatin to fluoresce. We prepared native chromatin from the cells using standard methods, treating isolated nuclei with micrococcal nuclease (MNase) and then extracting the soluble native chromatin using a high salt buffer. We next labeled the DNA within our chromatin preparations with TOTO-3, a red nucleic acid stain that is spectrally distinct from GFP. We wanted to analyze this dual labeled chromatin for two reasons. First, it permitted us to determine whether chromatin remained intact during nanofluidic electrokinetic flow, which is essential for any successful application of this method for epigenomic analysis. Intact chromatin will produce time-coincident TOTO-3 and GFP fluorescent SMD events that indicate DNA and histones were bound. Second, we wanted to demonstrate that simultaneous, multi-color detection of chromatin could be performed at high rates. This is a requirement if our platform is to be used to identify multiple epigenetic marks simultaneously on intact chromatin, on a genome wide basis.

Figure 2 illustrates a 0.25 s SCAN using chromatin extracted after a 5 min MNase digestion and then driven through a nanofluidic channel at 50 V. The top panel shows photon bursts corresponding to TOTO-3, which marks the DNA, while the bottom panel shows GFP fluorescence that marks the H2B. In order to identify the single molecule peaks shown (Figure 2), successive photon arrival times differing by less than 200 μ s were grouped as a burst. This time is about one-third that required for a molecule to pass through the inspection volume. Bursts with a total of 10 or more photons were designated as SMD events. The mean noise detected with buffer only was 0.22 photon/50 μ s. A velocity of 2.2 mm/s was calculated by fitting a histogram of the time duration (Figure S-1) for TOTO-3 SMD events. Since the sample solution was loaded into the reservoir connected to the negative electrode, only molecules carrying a net negative charge will be driven into the channel. In the pH 8.0 buffer used, only histone-free DNA and intact chromatin carry net negative charge. The net positive charge of individual free histones, each greater than 16.5 at this pH (see Supporting Information), will cause them to remain in the loading reservoir. The time-coincident SMD events are marked by blue bars spanning the plots, the majority of which are intact chromatin. Accurate identification of dual labeled chromatin was ensured by using optical filters that achieved more than 20 dB of spectral isolation of green GFP bursts from red TOTO-3 bursts, essentially eliminating fluorescence spectrum cross-talk or bleed-through during detection.

Approximately, ninety-three percent of all GFP molecules identified were coincident with a DNA molecule, indicating chromatin remained intact during electrokinetic flow within our nanofluidic device. This degree of coincidence was achieved using TOTO-3 at a 1:5 dye to base pair ratio, which provided the highest fluorescence intensity while not dissociating histones from the chromatin. Additional measurements (Table S-1) performed with 1:10 and 1:15 labeling ratios, produced similar levels of coincidence, further demonstrating that this labeling did not dissociate histones^{25, 26}. We observed that approximately one-third of the TOTO-3 signals were not coincident with GFP. This population of molecules is expected to arise from undigested and histone-free linker DNA known to exist in the genome, and also from intact chromatin that contained only the endogenous unlabeled H2B; half of the H2B in the cells is made from the GFP fusion transgene²⁴.

To further test the authenticity of chromatin detection using SCAN, we prepared chromatin from mixtures of nuclei from wild-type HeLa cells and HeLa cells with the H2B-GFP transgene, maintaining a constant total cell quantity for all mixtures. We anticipated that the rate of coincidence should drop with increasing amounts of wild-type chromatin, yielding fewer fragments with H2B-GFP. Fragments from each mixture were then detected by SCAN to determine the fraction of two-color labeled chromatin present in each mixture. We compiled (Figure 3a) a record of all coincident SMD events observed during a period of 15 min using a time coincidence histogram (TCH). The TCH contains the time-offset between all identified GFP events and TOTO-3 events within a fixed time window. The area under the peak and above the background level, describes the total chromatin fragments with H2B-GFP, which increased with the proportion of GFP-HeLa nuclei. We found the proportion of chromatin fragments with H2B-GFP to exhibit a linear increase with GFP-HeLa proportion for both the 5 and 15 min digestion series (Figure 3b). The direct proportion between chromatin fragments with H2B-GFP and input of GFP-HeLa nuclei validates the authentic detection of chromatin. Note that the slopes differ for the two digestion times with the prolonged digestion producing a smaller slope. The coincidence per DNA was principally reduced due to an increase in mononucleosome fragments and linker DNA fragments produced at longer digestion times, for a constant number of H2B-GFP nuclei within a mixture. This trend was also consistent with other chromatin preparations, wherein longer fragments demonstrated a higher coincidence proportion.

We observed variation in the size of chromatin fragments prepared during different batches of MNase digestion. For example, GFP-HeLa chromatin extracted after a 5 min MNase digestion, but prepared from different batches of GFP-HeLa cells on separate dates yielded different fragment sizes, as verified by gel electrophoresis (Figure S-3). We attribute this to variation in MNase activity. As a result, nanofluidic SCAN detected a 93% coincidence per GFP for samples with a median fragment size above 2 kbp (5 min lane in Figure S-3a); while 50% coincidence per GFP was observed for samples with a median fragment size less than 2 kbp (100% GFP, 5 min lane in Figure S-3b). To eliminate the possibility of systematic variation in the optical setup and/or nanofluidics, we later repeated these SCANS and obtained the same results. Further SCAN with chromatin digested to less than 1 kbp fragments (100% GFP, 15 min lane in Figure S-3b) demonstrated 35% coincidence per GFP, suggesting a trend that decreased nucleosome fragment size is correlated to decreased coincidence per GFP. We hypothesized that shorter nucleosome fragments, especially tightly-wrapped mononucleosome fragments, are less-likely to be identified in a coincident event due to decreased accessibility of TOTO-3 for nucleosomal DNA and an increasing amount of nucleosome free linker DNA that can appear with longer digestion times. This issue may be mitigated using a different fluorescent labeling method, extending SCAN to the study of short nucleosome fragments.

To evaluate the throughput of our nanofluidic device (Table 1), we utilized SCAN measurements of the 5 and 15 min digestions of 100% GFP-HeLa nuclei. The rates for SMD of both DNA and H2B were averaged over 15 min of analysis. We calculated the detection rate for time-coincident molecules of dual-labeled chromatin by performing a background-corrected TCH analysis for each minute of SCAN. The throughput for all molecule types was consistently higher for chromatin from the 5 min digest, as compared to the 15 min digest, and suggested a higher sample concentration. We verified this observation in-situ by fitting a histogram of the time between each SMD event of a given color with an exponential model to calculate the sample concentration¹⁶ (Figure S-2). The average fragment size was estimated separately using gel electrophoresis (Figure S-3) and then combined with the DNA detection rate to derive the analysis throughput.

Detection of DNA Methylation

Our next goal was to determine if probes that were successfully applied to epigenomic analysis in the past could be used to detect a *bona fide* epigenetic mark on our nanofluidic platform. We focused on DNA methylation and used MBD1 as our probe, which has been shown to bind methylated DNA specifically²⁷. Our test material was *Hind*III digested lambda DNA from a methylation deficient host, which we left unmethylated, or methylated *in vitro* using *Sss*I DNA methyltransferase. *Sss*I can methylate all 3,113 CpGs in the 48.5 kbp genome. We verified the effectiveness of the methylation reactions by digesting the DNA with the methylation sensitive restriction enzyme *Hpa*I (Figure S-4). Both DNA samples were stained with TOTO-3 and incubated with MBD1, which we labeled with Alexa Fluor 488. Alexa Fluor 488 is spectrally similar to GFP. The Alexa Fluor 488 labeled MBD1 retained its specificity for methylated DNA (Figure S-5). To facilitate optimum binding to methylated DNA, we added a molar excess of MBD1 to the stained DNA. We found that dilution of this mixture into our 1x TBE-based buffer resulted in stable electrokinetic flow, as indicated by consistent SMD rates and low-levels of non-specific interaction between probes and the nanofluidic structure.

Figure 4a illustrates the number of coincident SMD events for MBD1 mixed with unmethylated DNA (top panel) or methylated DNA (bottom panel). Each mixture was analyzed for 15 min at applied potential of 100 V. The background level of coincidence events in the unmethylated DNA sample was due to an excess of probe present in the mixture. However, the central Gaussian peak in the methylated DNA sample showed that bound MBD-DNA complexes were

detected above the high background. Methods have been described that could reduce background signals from free fluorescent probes ²⁸.

Similar to the chromatin analysis, we wanted to verify the authenticity of these detection events, so we prepared a dilution series of methylated and unmethylated DNA. We expected that with diminishing amounts of methylated DNA, we should observe a diminishing frequency of coincident events. This is consistent with our observed results (Figure 4b). There was a linear increase in the number of MBD-DNA complexes with increasing methylated DNA concentration, verifying the specificity of the signals detected by MBD1, and demonstrating the utility of nanofluidic SCAN for detecting *bona fide* epigenetic marks on individual molecules. Other fluorescence based approaches have been used to quantify DNA methylation ²⁹, but these were not single molecule methods.

CONCLUSIONS

We have described the development of SCAN using a nanofluidic platform to analyze individual molecules using the same fluorescently labeled probes that have been used in bulk epigenetic analysis in molecular biology. Confinement using nanofluidic channels enabled single molecule analysis to be performed within the 100-1000 pM concentration range, which was essential for maintaining chromatin structure^{30, 31}. We verified through multi-color SCAN that core histone octamers remained bound as nucleosomes within this nanofluidic environment during electrokinetic flow. Dilution of wild-type HeLa and GFP-HeLa nuclei was used to confirm the authenticity of coincidence detection. We observed chromatin fragments with a throughput of about 10 Mbp per minute using a single fluidic channel, indicating we could SCAN the entire genome of a fungal model organism in as few as eight minutes. We envision scaling our current throughput using parallel arrays with tens or hundreds of fluidic channels to perform analysis of larger organisms. For example, using 10 fluidic channels would allow 1x coverage of a 3 Gbp human genome to be scanned in just 30 minutes.

Further reduction of the nanofluidic channel cross-section would enable SMD at higher concentrations and increase the signal to noise ratio for single molecule fluorescent analysis. Since the probability of detecting a coincident event randomly is related to the sample concentration in the inspection volume, this probability can be engineered by reducing the channel cross-section. With sufficiently narrow channels the flowing chromatin or DNA molecules can be elongated allowing multiple fluorescent labels to be spatially separated and resolved. This may permit molecular mapping with spatial resolution sufficient to identify multiple epigenetic marks on a single nucleosome or to distinguish marks on adjacent nucleosomes. Prior work has shown optical resolution of molecular length to 114nm, equivalent to about 335bp, during rapid flow of lambda DNA³². This spatial resolution was possible using nanofluidic channels with cross sectional areas on the order of 0.01 μm^2 , an order of magnitude less than was used in these experiments. It is likely that smaller channels will allow us to identify molecules with multiple bound probes and resolve their positions on a chromatin fragment.

The applications of SCAN we described so far are analytical in nature, capable of revealing the frequency and coincidence of epigenetic mark placement in the genome. However, just as flow cytometers can have analytical or preparative capabilities, we believe the SCAN platform can be modified to rapidly sort chromatin fragments in real time based on their fluorescence signatures. This could provide an alternative to ChIP by allowing the selection and recovery of individual chromatin fragments with a desired set of epigenetic features. To identify exact epigenetic mark placement in a genome-wide context, SCAN sorted molecules could then be recovered for further analysis using PCR or DNA sequencing. Such an instrument would make SCAN-seq possible, an alternative to ChIP-seq. Importantly SCAN can monitor multiple

fluorescent marks simultaneously, enabling multiplexed, genome-wide epigenomic analysis not currently possible with ChIP.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge Adrian Bird for the MBD1 expression vector; Geoff Wahl for the H2B-GFP HeLa cells; Anders Lindroth, Madhukar Varshney, Cynthia Kinsland, and Jose Moran-Mirabal for helpful discussions. Funding was from the National Institute of Health (DA025722), Cornell Center for Vertebrate Genomics, Tang Family Scholars Program, and Cornell Nanobiotechnology Center. Fabrication of nanofluidic channels was conducted at the Cornell NanoScale Facility, a member of the National Nanotechnology Infrastructure Network, which is supported by the National Science Foundation (Grant ECS-0335765). This material is based upon work supported in part by the STC Program of the National Science Foundation under Agreement No. ECS-9876771.

REFERENCES

1. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. *Nature* 1997;389:251–260. [PubMed: 9305837]
2. Jenuwein T, Allis CD. *Science* 2001;293:1074–1080. [PubMed: 11498575]
3. Klose RJ, Bird AP. *Trends In Biochemical Sciences* 2006;31:89–97. [PubMed: 16403636]
4. Mikkelsen TS, Ku MC, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie XH, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE. *Nature* 2007;448:553–U552. [PubMed: 17603471]
5. Feinberg AP. *Nature* 2007;447:433–440. [PubMed: 17522677]
6. Waterland RA, Jirtle RL. *Molecular And Cellular Biology* 2003;23:5293–5300. [PubMed: 12861015]
7. Anway MD, Cupp AS, Uzumcu M, Skinner MK. *Science* 2005;308:1466–1469. [PubMed: 15933200]
8. Weaver ICG, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ. *Nature Neuroscience* 2004;7:847–854.
9. Bernstein BE, Meissner A, Lander ES. *Cell* 2007;128:669–681. [PubMed: 17320505]
10. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, Volkert TL, Wilson CJ, Bell SP, Young RA. *Science* 2000;290:2306–+. [PubMed: 11125145]
11. Barski A, Cuddapah S, Cui KR, Roh TY, Schones DE, Wang ZB, Wei G, Chepelev I, Zhao KJ. *Cell* 2007;129:823–837. [PubMed: 17512414]
12. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D. *Nature Genetics* 2005;37:853–862. [PubMed: 16007088]
13. Zhang XY, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen HM, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR. *Cell* 2006;126:1189–1201. [PubMed: 16949657]
14. O'Neill LP, VerMilyea MD, Turner BM. *Nature Genetics* 2006;38:835–841. [PubMed: 16767102]
15. Bernstein BE, Mikkelsen TS, Xie XH, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. *Cell* 2006;125:315–326. [PubMed: 16630819]
16. Foquet M, Korlach J, Zipfel W, Webb WW, Craighead HG. *Analytical Chemistry* 2002;74:1415–1422. [PubMed: 11922312]
17. Neely LA, Patel S, Garver J, Gallo M, Hackett M, McLaughlin S, Nadel M, Harris J, Gullans S, Rooke J. *Nature Methods* 2006;3:41–46. [PubMed: 16369552]
18. Nolan RL, Cai H, Nolan JP, Goodwin PM. *Analytical Chemistry* 2003;75:6236–6243. [PubMed: 14616007]
19. Stavits SM, Corgie SC, Cipriany BR, Craighead HG. *Biomicrofluidics* 2007;1
20. Kornberg RD, Lapointe JW, Lorch Y. *Methods In Enzymology* 1989;170:3–14. [PubMed: 2770543]

21. Foquet M, Korlach J, Zipfel WR, Webb WW, Craighead HG. *Analytical Chemistry* 2004;76:1618–1626. [PubMed: 15018559]
22. Stavis SM, Edel JB, Samiee KT, Craighead HG. *Lab On A Chip* 2005;5:337–343. [PubMed: 15726210]
23. Ambrose WP, Goodwin PM, Jett JH, Van Orden A, Werner JH, Keller RA. *Chemical Reviews* 1999;99:2929–2956. [PubMed: 11749506]
24. Kanda T, Sullivan KF, Wahl GM. *Current Biology* 1998;8:377–385. [PubMed: 9545195]
25. Wojcik K, Dobrucki JW. *Cytometry Part A* 2008;73A:555–562.
26. Martin RM, Leonhardt H, Cardoso MC. *Cytometry Part A* 2005;67A:45–52.
27. Jorgensen HF, Adie K, Chaubert P, Bird AP. *Nucleic Acids Research* 2006;34
28. Wang X, Song Y, Song M, Wang Z, Li T, Wang H. *Anal Chem* 2009;81:7885–7891. [PubMed: 19788313]
29. Stach D, Schmitz OJ, Stilgenbauer S, Benner A, Dohner H, Wiessler M, Lyko F. *Nucleic Acids Res* 2003;31:E2. [PubMed: 12527791]
30. Hagerman TA, Fu Q, Molinie B, Denvir J, Lindsay S, Georgel PT. *Biophysical Journal* 2009;96:1944–1951. [PubMed: 19254554]
31. Claudet C, Angelov D, Bouvet P, Dimitrov S, Bednar J. *Journal Of Biological Chemistry* 2005;280:19958–19965. [PubMed: 15772075]
32. Reccius CH, Stavis SM, Mannion JT, Walker LP, Craighead HG. *Biophysical Journal* 2008;95:273–286. [PubMed: 18339746]

Figure 1.

Experimental platform. (A) A differential interference contrast optical micrograph of a typical nanofluidic channel used in SCAN. The narrow region, with a 500 nm wide and 250 nm deep cross-section, was used during fluorescence detection. We formed 432 of these channels on a single 100 mm diameter fused silica wafer. The scale bar is 10 μm . (B) Schematic diagram of a wafer mounted on a confocal fluorescence microscope. Two overlapped lasers illuminated a 1.3 μm length of the nanofluidic channel and formed an inspection volume of 0.16 fL. Collection of the dim, fluorescent signature for each molecule was achieved using a confocal aperture, which spatially restricted the optical collection to the inspection volume, and avalanche photodiodes (APDs), which provided single photon detection.

Figure 2.

Process of Single Molecule Detection and Two-Color Coincidence Analysis. Time-trace record of photon bursts observed by each APD, showing 0.25 seconds of a 15 minute nanofluidic SCAN. A burst with a sum of 10 or more photons satisfied a threshold condition and was designated a SMD event, shown here by a red or green marker identifying DNA and histone H2B respectively. Intact chromatin fragments, highlighted in blue, were identified by time-coincident detection of both a red and green event.

Figure 3.

Nanofluidic SCAN of GFP-HeLa Chromatin at Different Digestion Times. Chromatin was isolated from the nuclei of wild-type HeLa cells and HeLa cells with a H2B-GFP fusion transgene and then analyzed by SCAN for 15 minutes. (A) A TCH illustrates the absence of coincident two-color SMD events when analyzing chromatin from wild-type HeLa nuclei. With GFP-HeLa chromatin from the 5 minute digestion, a central Gaussian peak, corresponding to intact chromatin molecules emitting two fluorescent colors, emerged from a background of uncorrelated events. By integrating the area under the peak and subtracting the uncorrelated background, we observed more than 16,000 two-color chromatin molecules. (B) The proportion of two-color chromatin molecules increased in direct proportion with GFP-HeLa nuclei content, as described by a linear fit with $R^2=0.98$ and $R^2=0.95$ for the 5 and 15 minute digestion assays, respectively. Error bars represent the propagated error from SMD of both the bound and unbound molecules.

Figure 4.

Detection of DNA Methylation. (A) Unmethylated (top) and methylated (bottom) DNA samples labeled with TOTO-3 were both incubated with a molar excess of MBD1 probes labeled with Alexa Fluor 488 and then analyzed for 15 minutes. The emergent peak in the bottom panel demonstrates SMD of methylated DNA. The molar excess of labeled MBD1 contributed to the background of uncorrelated events within each experiment. (B) We analyzed mixtures of methylated and unmethylated DNA. The proportion of dual-color labeled MBD-DNA was shown to increase with methylated DNA, as described by a linear fit with $R^2=0.99$. Error bars represent the propagated error from SMD of both the bound and unbound molecules.

Table 1

Nanofluidic SCAN Throughput for 100% GFP-HeLa Samples

MNase Treatment (min)	5	15
DNA Concentration (pM) ^b	588 ± 8	248 ± 2
Dual-Labeled Chromatin (molecules/min) ^a	1067 ± 114	201 ± 36
H2B (molecules/min) ^a	2116 ± 143	568 ± 67
DNA (molecules/min) ^a	6238 ± 611	2287 ± 184
Gel Estimated Average Fragment Size (bp) ^c	1600	500
DNA Throughput (Mbp/min) ^d	10	1

^a Analysis for a 15 minute SCAN at an electrokinetic drive of 50 Volts.^b DNA concentration measured in-situ based upon the time between SMD events^c Fragment size estimated by gel electrophoresis^d DNA detection rate and estimated fragment size are used to calculate DNA throughput.