

Rapid, solid-phase based automated analysis of chromatin structure and transcription factor occupancy in living eukaryotic cells

Richard Ingram, Hiromi Tagoh, Arthur D. Riggs¹ and Constanze Bonifer*

Molecular Medicine Unit, University of Leeds, St James's University Hospital, Leeds LS9 7TF, UK and
¹Division of Biology, Beckman Institute of City of Hope, 1500 Duarte Road, Duarte, CA 91010, USA

Received October 13, 2004; Revised and Accepted December 3, 2004

ABSTRACT

Transcription factors, chromatin components and chromatin modification activities are involved in many diseases including cancer. However, the means by which alterations in these factors influence the epigenotype of specific cell types is poorly understood. One problem that limits progress is that regulatory regions of eukaryotic genes sometimes extend over large regions of DNA. To improve chromatin structure–function analysis over such large regions, we have developed an automated, relatively simple procedure that uses magnetic beads and a capillary sequencer for ligation-mediated-PCR (LM-PCR). We show that the procedure can be used for the rapid examination of chromatin fine-structure, nucleosome positioning as well as changes in transcription factor binding-site occupancy during cellular differentiation.

INTRODUCTION

After the completion of the sequencing of the genome of many model organisms, one of the major challenges in modern biology is the elucidation of how regulatory information is processed at the chromatin or epigenetic level. At individual genes, epigenetic regulatory proteins influence the transcriptional status and the accessibility to the transcription machinery by modifying chromatin structure (1,2). In addition, chromatin architecture has a profound impact on transcription factor–chromatin interactions. Not all transcription factors can interact with nucleosomal DNA, and some factors actively position nucleosomes, thus creating precisely arranged substrates for chromatin modification activities (3,4). With respect to the nucleosome position, a shift of a small number of base-pairs can greatly influence factor binding, and factors bind most readily to nucleosome-free regions (5). Such

alterations in the chromatin structure and the transcriptional control of specific genes underlie all developmental processes. Moreover, it is now abundantly clear that aberrant regulation of key genes at the epigenetic level can cause the deregulation of normal differentiation processes. For example, in leukaemia it is often the expression of one aberrant transcription factor that triggers a cascade of events starting with the recruitment of inappropriate chromatin modification activities to the cis-regulatory elements of specific genes and their subsequent deregulation (6). This can have severe consequences for a specific differentiation pathway that depends on the coordinated activation and silencing of specific genetic programs. To understand all aspects of the impact of normal and aberrant transcription factors on chromatin architecture, it is of vital importance to develop rapid high-resolution chromatin structure analysis methods. To this end, we developed a rapid, sensitive and highly versatile automated procedure that will greatly facilitate the analysis of different chromatin features in living eukaryotic cells.

MATERIALS AND METHODS

Cell culture

NIH3T3 and RAW264 cells were grown in DMEM supplemented with 10% foetal calf serum (FCS), 100 U/ml penicillin and 100 mg/ml streptomycin.

In vivo footprinting and automated single-strand break specific LM-PCR on the *c-fms* promoter

Cultured cells were treated with 0.2% dimethyl sulfate (DMS) in PBS before DNA extraction and piperidine treatment as described in (7). To perform automated ligation-mediated PCR (LM-PCR), ~1 µg DNA, previously treated with DMS and piperidine, in 2 µl dH₂O or 0.1× TE was placed in the wells of a 96-well plate which was then positioned on a Biomek 2000 robotic workstation in the pre-chilled 96-well plate holder (B6; see Figure 1). A detailed description of the

*To whom correspondence should be addressed. Tel: +44 113 206 5676; Fax: +44 113 244 4475; Email: c.bonifer@leeds.ac.uk

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use permissions, please contact journals.permissions@oupjournals.org.

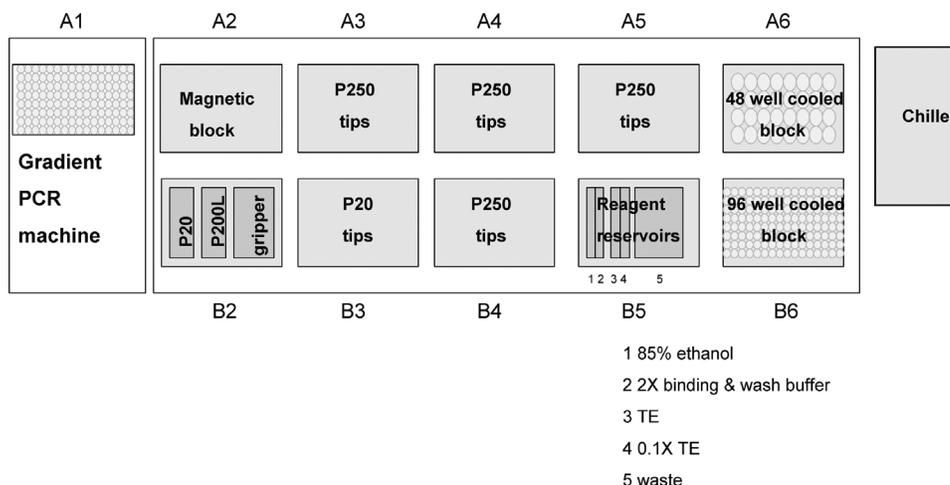


Figure 1. A schematic representation of the layout of the components of the robotic workstation.

set-up of the robotic workstation is shown in Figure 1, and the exact settings for the robotic workstation are detailed in Supplementary Material. Primers were as in (8).

Primer extension. An aliquot of 3 μ l of the primer extension master mixture was added to give a final concentration of 1 \times ThermoPol buffer (NEB), 250 μ M dNTPs, 1 pmol biotinylated first primer, 2 U Vent polymerase Exo- (NEB), 5% DMSO, and mixed five times by pipetting. The reaction was then overlaid with 15 μ l of mineral oil. The reaction plate was transferred to the preheated PCR machine (95°C with a lid heated to 105°C). PCR program: 95°C 10 min, 52°C for 15 min, 72°C for 20 min, 10°C for 3 min.

Ligation. To reduce the time needed for the ligation, PEG 6000 was added to the ligation reaction at 10% final concentration to concentrate the DNA molecules. In addition, during the ligation reaction, temperature was cycled between 16°C, which allowed the DNA molecules to come together while still retaining some ligase activity, and 37°C where the ligase is most active. The PEG 6000 containing buffer could be successfully pipetted by setting a very low pipetting speed. After the primer extension reaction, the plate was moved to the 96-well cool position (B6) and 7.21 μ l of ligation master mixture was added to give a final concentration of 29 mM Tris-HCl (pH 7.5), 8.0 mM MgCl₂, 20 mM DTT, 1.0 mM ATP, 0.05 mM BSA, 10% PEG 6000 40 pmol LP25-21 linker [linker made as described previously (7), and 2.88 U T4 Ligase (Promega)]. This was mixed very slowly by pipetting five times before returning to the PCR machine. PCR program: 16°C for 20 min, 37°C for 10 s for 16 cycles.

Ligation product capture. Dynal M270 streptavidin coated beads (150 μ g) in 12.5 μ l of 2 \times BW buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2.0 M NaCl) were added to each sample and incubated at room temperature for 60 min with mixing every 20 min by pipetting up and down five times. The plate was then moved to the magnetic separator (A2) and the system was paused for 150 s, the liquid was slowly removed from the beads and replaced with 50 μ l 2 \times BW buffer, mixing vigorously five times to resuspend the beads. System was paused for 150 s for separation, 2 \times BW buffer was removed

slowly and replaced with 50 μ l TE (10 mM Tris, pH 7.5, 1 mM EDTA) with mixing as described before. The system was paused for separation and the TE was slowly removed. The TE wash was then repeated. The success of the amplification step is dependent on the retention of the beads with the captured target DNA on the plate after washing. In our system, the beads are drawn to the side of the tube by the magnet. Hence, careful positioning of the tip at the very bottom of the well combined with slow pipetting speeds when drawing off the buffer is important for retaining the magnetic beads with the captured template for the following PCR reaction.

PCR-amplification of ligation products. The beads were resuspended in 10 μ l 0.1 \times TE overlaid with 15 μ l mineral oil and the plate transferred to the PCR machine to denature the beads at 95°C for 15 min. The plate was moved back to cool position and 40 μ l of amplification master mix to give a final concentration of 1 \times cloned Pfu buffer (Stratagene), 1.4 M Betaine, 250 μ M dNTPs, 10 pmol LP25 primer, 10 pmol 2nd gene-specific primer; 2.5 U Pfu turbo DNA polymerase (Stratagene) was added, before the plate was returned to the pre-heated PCR machine. PCR program: 95°C for 5 min (95°C for 45 s, 60°C for 3 min, 72°C for 5 min for 22 cycles), 72°C for 10 min, heated lid 105°C.

Labelling reaction. An aliquot of 5 μ l of the amplification reaction was transferred to a new plate to which 4 μ l of labelling mix was added to a final concentration of 1 \times Pfu buffer, 1.25 M Betaine, 250 μ M dNTPs, 0.5 U Pfu turbo DNA polymerase, 1 pmol fluorescent D4 labelled third primer (Beckman Coulter, ProliGo). The mix was overlaid with 15 μ l of mineral oil. The labelling plate was then transferred to the PCR machine. PCR program: 95°C for 5 min (95°C for 45 s, 60°C for 3 min, 72°C for 5 min for 7 cycles), 72°C for 10 min, 8°C for 5 min (further details of the running program are described in Supplementary Material). Primer sequences and radioactive labelling of samples are described previously (8).

Automated double-strand break LM-PCR

Micrococcal nuclease (MNase) digested and linker ligated genomic DNA was prepared exactly as described in (9).

Approximately 0.1–0.2 μg MNase digested and linker ligated DNA in 5 μl dH_2O was placed into a 96-well plate, which was then positioned on the Biomek 2000 workstation into a pre-chilled 96-well plate holder.

Primer extension reaction. An aliquot of 25 μl of the primer extension master mix was added to give a final concentration of $1\times$ ThermoPol buffer (NEB), [250 μM dNTPs, 2 pmol biotinylated first primer, 2 U (Exo-)Vent DNA polymerase (NEB), 5% DMSO], and was mixed five times by pipetting. The reactions were then overlaid with 15 μl of mineral oil. The reaction plate was transferred to the PCR machine preheated to 95°C with heated lid (105°C) for program: 95°C for 5 min, 52°C for 10 min, 72°C for 10 min (95°C for 45 s, 52°C for 3 min, 72°C for 10 min for 11 cycles or less, dependent on the required sensitivity), 10°C for 3 min. Dynal M270 streptavidine coated beads (150 μg) in 30 μl of $2\times$ BW buffer were added to each sample and incubated at room temperature for 60 min with mixing at every 20 min by pipetting up and down for five times. Bead washing, amplification and labelling were as for single-stranded LM-PCR. Further details of the running program are described in Supplementary Material.

Primer design and LM-PCR primer optimization

OligoTM5.1 computer program was used to design the primers. The lengths of the primers should normally be between 21 and 25 bases and the T_d of successively used primers should increase in the order $1 < 2 < 3$. Primers should be nested but it is not important whether the 1st and 2nd primers overlap or not. However, the 2nd and 3rd primers should ideally overlap as this avoids interference between the primers.

The primer extension reaction was set up manually across one row of a 96-well plate and manually placed into the PCR machine. The primer extension program was entered manually into the gradient PCR machine with the annealing temperature gradient centred around the predicted annealing temperature ($T_m - 3^\circ\text{C}$) of the primer with a 12°C span. The PCR reaction and the Biomek program were started simultaneously, but the workstation was paused for 60 min to allow the primer extension program to finish. After 60 min, the robot opened the PCR machine lid and continued. Further experimental details are described in Supplementary Material. In the labelling reaction, the D4 labelled specific third primer was replaced by D4 labelled LP25.

Removing excess primers for sample analysis on a capillary sequencer

To each sample, 10 μl of CleanSeq (Agencourt Bioscience) and 40 μl of 85% ethanol was added and mixed well by pipetting seven times. Samples were incubated for 60 s, then transferred to magnet for 5 min separation. All liquid was removed from the beads and replaced with 100 μl 85% ethanol without mixing and incubated for 30 s. The ethanol was then removed and the samples were air dried for 10 min before being removed from the magnet. The beads were then resuspended in 40 μl of sample loading solution (SLS) (Beckman Coulter) containing 0.1 μl of 400 bp marker (Beckman Coulter), and incubated at room temperature for 5 min. The plate was returned to the magnet for separation (5 min) and 35 μl of SLS were transferred to a clean well. The

plate was stored in the dark until the samples could be loaded onto the capillary sequencer.

Analysis of samples on a capillary sequencer

Cleaned up samples are transferred manually to a CEQ sample plate and overlaid with mineral oil. The beads in the SLS need to be removed before loading onto a CEQ 8000 sequencer (Beckman Coulter) as this results in sharper peaks and longer traces. It is common practice to heat denature sequence reactions before they are run on a gel/sequencer, however, we found that sufficient denaturation is provided by the formamide in the loading solution. Further denaturation by heating reduces the quality of the signal substantially, hence this option should be turned off on the sequencer. In summary, running parameters were as follows: Injection—2.5 kV 30 s; Denaturation—off; Capillary temperature—45°C; Wait for capillary temperature—Yes; Separation—6 kV for 35 min. Analysis parameters: Slope threshold—8%; Relative peak height threshold—2%. All other settings were left at default. These analysis parameters may have to be varied if clear peaks are not recognized.

Peak height analysis

Using the Peak Analysis option within the CEQ8000 software, all the peaks of interest were selected and the grid exported to Excel. An average fluorescence for each trace was calculated, by exporting the fragment list to Excel and taking an average of all peaks above 100 bases. The peak height readings for the previously selected peaks were divided by that average to produce normalized peak height readings. The normalized peak height reading was then divided by the normalized corresponding peak from the G reaction to produce a peak height ratio relative to naked DNA.

RESULTS AND DISCUSSION

In vivo genomic footprinting is a powerful method to investigate different features of eukaryotic chromatin at nucleotide resolution (10,11). It consists of two steps, the first of which is to create DNA strand breaks by chemical or enzymatic DNA cleavage, and the second step is to make these lesions visible. A number of different DNA modifying agents can be used for chromatin structure studies. This includes treating cells with DMS, KMnO_4 or UV irradiation, or digesting nuclei with nucleases such as restriction enzymes, MNase and DNaseI. As the action of such agents is modified by chromosomal proteins, it is possible to draw conclusions about *in vivo* chromatin fine-structure by comparing the position and frequency of DNA modifications carried out in living cells to naked DNA modified *in vitro*.

DNA lesions generated by the different chromatin modifying agents described above are visualized by ligation-mediated-PCR (LM-PCR), which usually starts with a primer extension reaction using a gene-specific primer and genomic DNA as template (11). Use of biotinylated primers and magnetic beads to capture primer extension products improves sensitivity and quality (7,8,12). We have further adapted this method for the selective amplification of double-strand cuts (9). A schematic outline of both amplification procedures is depicted in Figure 2. LM-PCR products can be analysed by

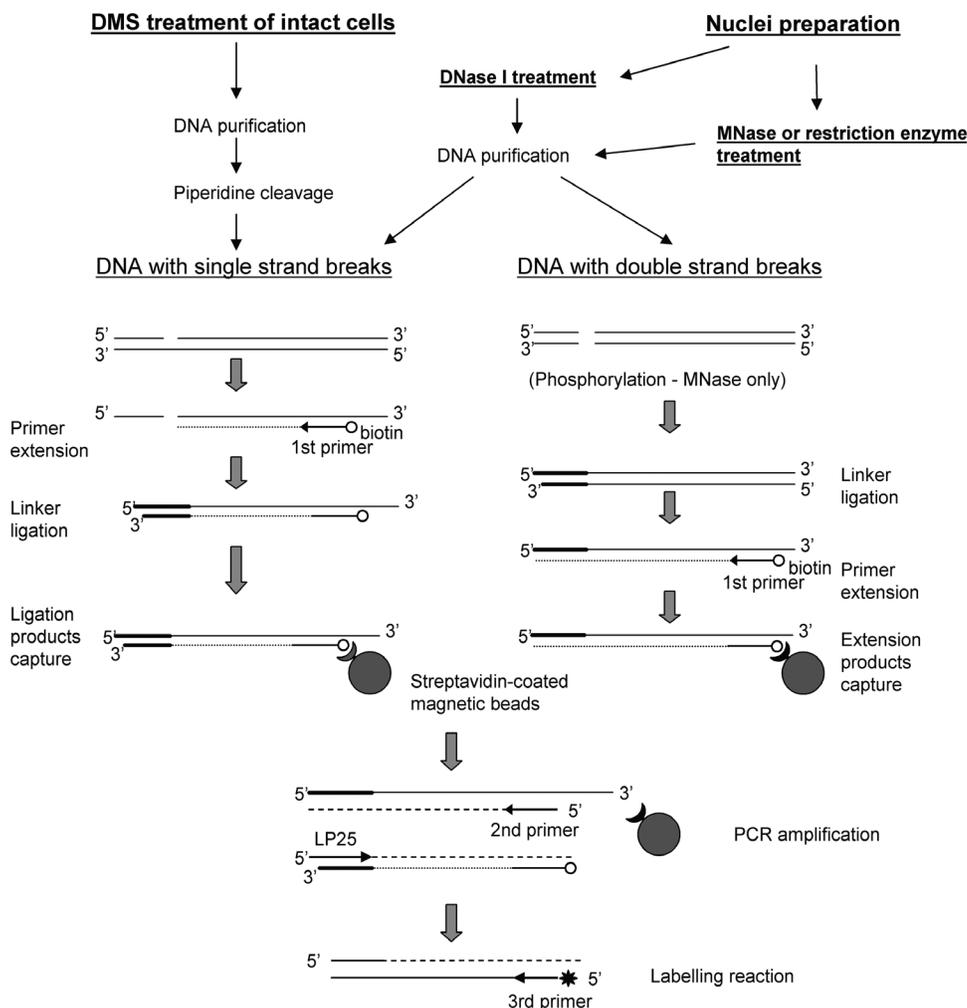


Figure 2. Flowchart of different LM-PCR methods for the examination of transcription factor occupancy and chromatin structure.

either labelling them with radioactive or fluorescent primers (7,13,14). Using these methods, we can gather information about (i) the precise position of transcription factors interacting with DNA by DMS and DNaseI treatment, (ii) translational and rotational nucleosome positioning by mapping sites of MNase and hydroxyl radical cleavage or (iii) differences in chromatin accessibility to DNaseI and restriction enzyme digestion.

One advantage of the automated method reported here as compared to earlier described procedures (14,15) is that all genomic DNA template is removed after extension-product capture. This effectively eliminates most background problems caused by non-specific priming in the PCR amplification reaction following the primer extension step. However, all reported methods have in common that they are highly dependent on the quality of the first extension primer. To conduct reproducible experiments, it is therefore vital to perform a series of temperature optimization tests (see below). To perform primer optimization manually takes ~3–4 days from the first amplification reaction to the analysis of bands on a gel, thus primer optimization and the analysis of even one cis-regulatory element can take many weeks. To shorten the procedure, to make it more reproducible and to increase sample throughput, we adapted these methods to be performed on a

robotic workstation. In addition, we developed alternative sample analysis methods by running samples on a capillary sequencer with a laser detection system capable of detecting near-infrared signals with high sensitivity (in our case, a CEQ8000). On this instrument, a run of eight samples using pre-made capillaries takes only 50 min and the peaks detected by the laser can be accurately quantified using appropriate software. Using the manual procedure, we have already shown that the solid-phase LM-PCR protocol is very sensitive and capable of detecting subtle footprints and chromatin fine-structure alterations (8,9,16) in rare precursor cells. The procedure reported here is of similar sensitivity and gives excellent, reproducible data showing both DMS footprints and, by use of MNase, information on nucleosome positioning. Each of these experiments will be described in more detail after primer optimization is discussed.

Automated primer optimization

One of the major advantages of the automated system is to be able to rapidly determine the optimal annealing temperature for the primer extension reaction. From our experience, it is not possible to reliably do this by conventional PCR, an actual LM-PCR experiment has to be performed. To this end, we

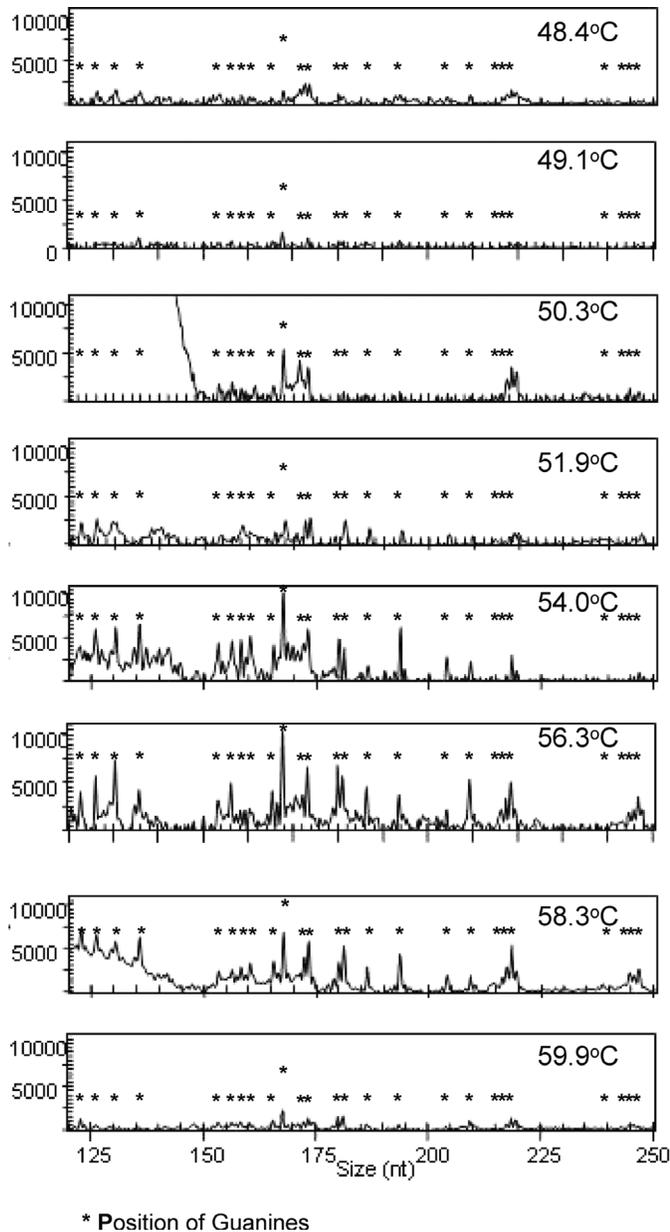


Figure 3. An example of automated primer optimization. CEQ 8000 traces from primer optimization of a new *c-fms* primer showing optimal results for 56°C. The first two traces show that the temperature is too low allowing the primer to anneal randomly giving virtually no signal. As the temperature increases to 50°C, the primer anneals to several places giving a number of different products, all of which are amplified by the labelled LP25 producing an undifferentiated mass of signal at the start of the trace. From 52°C through to 58°C, the primer binds specifically with the best result at 56°C, above this temperature the primer binds inefficiently resulting in loss of signal. The * mark the position of the Gs within the sequence.

have therefore devised an automated protocol that uses a gradient PCR machine and labelling with a fluorescent LP25 primer. If the primer annealing temperature is too low, mispriming will occur, which can result in both, higher background and lower amplification efficiency. An example for such a primer optimization experiment is shown in Figure 3. It shows that if labelling is carried out using a labelled LP25, the presence of non-specifically amplified DNA can also be

detected (an example for this phenomenon can be seen in Figure 3, third panel). Primer optimization is best carried out on DMS-treated genomic DNA as this will only reveal the presence of guanines (and to a lesser extent adenines) and background bands can easily be identified. The annealing temperature of the second primer is generally set at 3°C below the predicted T_m in 50 mM NaCl. This generally gives good results, but if the signal is still weak or contains significant background, optimization of the second primer can be carried out in an experiment using the best temperature for the first primer obtained in the initial primer optimization experiment. If after trying to optimize both primers, the results are unsatisfactory, then it is unlikely that the primer set will work and new primers should be designed.

If the first primer is superior, as in the example depicted in Figure 3, it is theoretically possible to directly use a labelled LP25 primer for sample analysis. However, even with good primers that give a clearly readable sequence, some non-specifically amplified DNA is often still present which will obscure signal quantification. Such background signals are efficiently eliminated by labelling with a third gene-specific primer.

Sample analysis

Removing unincorporated primers is essential to be able to analyse LM-PCR samples on a capillary sequencer as the amount of free primer far exceeds the amount of LM-PCR product. Excess primer and salts compete with the labelled product for capillary space thus reducing both signal strength and quality. Cleaning up of the labelling reactions with a primer-removal system such as CleanSeq beads removes free primers, all salts and other contaminants resulting in sharper and stronger traces. As CleanSeq is a magnetic bead system, this procedure was also incorporated into the automated procedure and carried out by the workstation (Supplementary Materials).

Peaks were quantified using the peak analysis software on the CEQ8000 sequencer. To normalize peak heights, we measured peak heights of all relevant peaks across a wide region of interest and use these to calculate an average for the total fluorescence of each trace. These normalized peak heights were then used as the basis for calculating peak height ratios at the guanines. If it is desired to generate a gel image, samples prepared using D4 labelled fluorescent primers can also be analysed on other sequencers able to detect near-infrared signals, such as LiCOR (14). To date only sequencers using near-infrared detection have given adequate sensitivity.

Single-strand-specific LM-PCR

For a number of years our laboratory has used the macrophage-colony-stimulating-factor receptor locus or *c-fms* as a model to study cell fate decisions at the epigenetic level (8,16,17). A map of this locus is depicted and the position of its cis-regulatory elements is given in Supplementary Figure 1. To confirm that the automated system produces reliable results, we performed single-stranded LM-PCR both manually and by automation on a region of the *c-fms* promoter, which contains a known DMS-footprint in a PU.1 binding site in macrophages. As template, we used DMS-treated genomic DNA as well as DNA from DMS-treated *c-fms* non-expressing

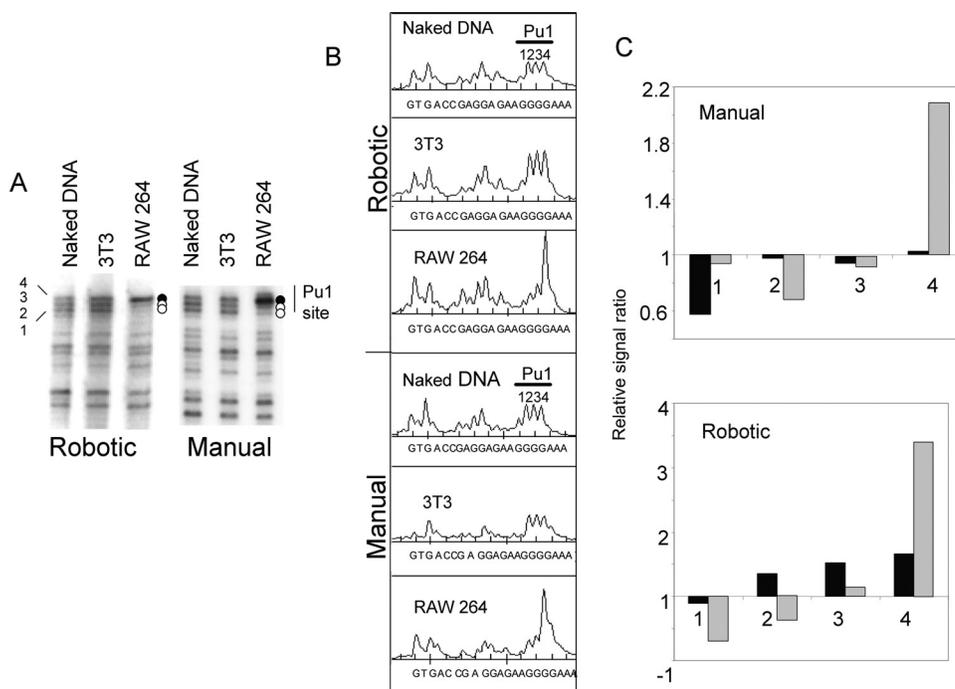


Figure 4. Comparison of manual and automated single-strand LM-PCR. (A) *In vivo* DMS footprinting followed by single-strand LM-PCR of a section of the murine *c-fms* promoter containing a PU.1 binding site, which is only occupied in macrophage cells (8). The same samples were run on a conventional radioactive sequencing acrylamide gel. The footprint is only visible in the RAW264 lanes, the protected bands are marked with open circles and the enhanced band by the filled circle. (B) The same samples were run on a CEQ 8000 capillary sequencer. The relevant sequence is depicted below the traces. The footprint can be seen as a reduction of peaks 2 and 3 and an increase of peak 4 in the RAW264 traces. CEQ 8000 dyes are excited to fluoresce using diode lasers. D4, which is excited at 685 nm is used as standard label and emits at 706 nm. (C) Graphs produced from CEQ 8000 peak height analysis showing the protection of peaks 2 and 3 and the DMS-hyperreactivity at peak 4 in RAW264 cells, in both manually and robotically produced samples.

NIH3T3 fibroblasts and from the *c-fms* expressing RAW264 macrophage cell line. Figure 4A shows radioactively labelled products from this experiment analysed on an acrylamide gel. It can be seen that the pattern of bands is comparable in both manual and automated samples and that the footprint in the PU.1 site is clearly visible in both RAW264 lanes. The RAW264 samples show a protection of guanines 2 and 3, marked by the open circles, and an enhancement of DMS reactivity at G4, marked by the filled circle when compared to 3T3 cells and naked DNA.

The same samples were run on the capillary sequencer (Figure 4B) where the footprint was also clearly visible. In RAW264 DNA, it can be seen that there is a reduction in peak height of Gs 2 and 3 and a big increase in the height of the G4 peak, in comparison to both 3T3 and DMS-treated genomic DNA. These results show that automated LM-PCR results are as good as those produced manually and that footprints can be identified clearly in CEQ 8000 traces. One of the advantages of running samples on the capillary sequencer is the ability to quantify the peak heights and hence produce peak height ratios relative to the naked DNA control. Figure 4C shows peak height ratios for the four guanine residues within the PU.1 binding site as compared to genomic DNA. Graphs for both traces show the same pattern of protection and enhancement of DMS reactivity at the four guanines in the PU.1 binding site. The footprinting pattern analysed this way was highly reproducible. There was only very little variability between experiments as demonstrated by quantifying results of six different independently performed footprinting reactions

(Supplementary Figure 1). Another example for a successful automated footprinting reaction analyzing the intronic enhancer of the *c-fms* gene (8) is depicted in Supplementary Figure 2. To date, we have applied this method to a number of cis-regulatory elements of different genes.

Double-strand break specific LM-PCR

Figure 5A shows automated LM-PCR on MNase-digested genomic DNA as well as DNA isolated from MNase-digested nuclei from mouse bone-marrow derived macrophages and embryonic fibroblasts (MEFs), again looking at the *c-fms* promoter. It is obvious that the digestion pattern differs strongly between samples. In macrophages, a cluster of MNase-hypersensitive sites are seen around position -67 bp upstream of the ATG. This is not seen in MEFs, instead there is an enhancement of digestion around -130 bp that is not seen in macrophages. The same pattern is clearly visible on a gel loaded with radiolabelled samples (Figure 5B). To control for equal digestion efficiency, primers spanning the GAPDH promoter were used in a separate automated LM-PCR experiment and run on a gel and the capillary sequencer (Figure 5C and D). In different experiments, we have shown that this region has a specifically positioned nucleosome that adopts a different position in non-macrophage cells (9). This example clearly demonstrates the feasibility of our automated system for double-strand specific LM-PCR. In principle, this method can also be applied to DNA-cleaved with restriction enzymes generating blunt ends. Restriction enzymes generating 5' or 3'

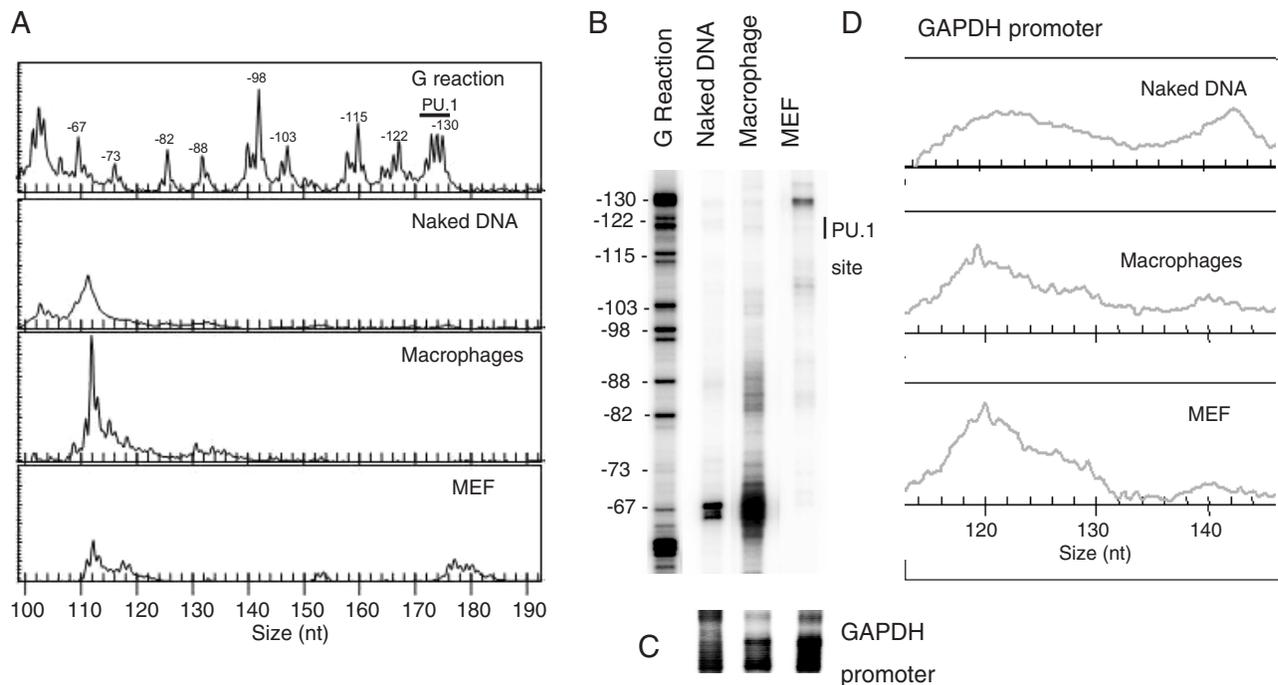


Figure 5. Double-strand-specific break automated LM-PCR. (A) Capillary sequencer traces showing differences in MNase accessibility at the *c-fms* promoter between macrophages and fibroblasts. The position of the PU1 binding site and the distance from the major peaks relative to the ATG are indicated. (B) The same, but radiolabelled samples run on an acrylamide sequencing gel. The bands in the G reaction corresponding to the major peaks in the capillary sequencer trace are marked along with the transcription start sites. (C) Gel showing no differences in the GAPDH promoter MNase digestion control. (D) CEQ traces on GAPDH promoter digestion control showing a very similar pattern between the samples. Primers for the internal GAPDH control are coupled to the D3 dye, which excites at 650 nm and emits at 670 nm.

overhangs can also be used to probe accessibility, but here a blunt end has to be generated first by removal or end-filling of any single-stranded sequence.

Outlook

With our present set-up, we can only process 18 samples in parallel, but this number could be increased by adding an extension to the Biomek 2000 workstation that allows to analyse 96 samples in parallel. Provided the annealing temperatures for each primer are similar, an extended system can theoretically process 8 different DNA samples with 12 independent primers in ~ 24 h. For the capillary sequencer, the usable read begins generally in the region of 20–30 bases beyond the labelling primer and extends for 200–250 bases, making it possible to examine the chromatin structure of 3 kb of regulatory sequence from one strand in up to eight different cell types. This is a vast improvement over the manual procedure.

The method we describe here is not only of immediate value for any group working on individual genes, such as the target genes for normal and oncogenic transcription factors, but also has the potential to contribute to more global strategies of gene expression analysis. A number of different efforts are now underway to identify functional elements in eukaryotic genes in a global fashion using a number of different high-throughput methods. This includes the identification of conserved regulatory elements using bioinformatics approaches, as well as global efforts to identify regions encompassing areas of remodelled chromatin such as DNaseI hypersensitive sites

that are likely to contain active cis-elements [summarized in (18)]. Such data will guide where to home in with our more detailed studies looking at individual genes. Once regulatory regions are identified, our automated method will enable the rapid examination of cis-regulatory elements of one gene in one cell type or the comparison of one cis-element between different cell types. Such experiments will not yield unambiguous information about the identity of transcription factors binding to a specific recognition sequence or the nature of histone modifications, but combined with other high-throughput methods such as chromatin immunoprecipitation–microarray (ChIP-on-chip) assays (19), they can provide exquisitely detailed information about the chromosomal organization and the transcription factor occupancy of active and inactive genes. However, currently our ability to handle large sample numbers and our ability to process very large amounts of data are limited. A remaining challenge is therefore to develop software and hardware tools allowing chromatin fine-structure analysis in a truly high-throughput fashion.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by the grants from the Leukaemia Research Fund, the Biological and Biotechnological Science Research Council (BBSRC), the City of Hope Medical Centre

and the Wellcome Trust to C.B. Hiromi Tagoh is a recipient of a Kay Kendall Leukaemia Fund Fellowship. Funding to pay the Open Access publication charges for this article was provided by the University of Leeds.

REFERENCES

- Narlikar,G.J., Fan,H.Y. and Kingston,R.E. (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, **108**, 475–487.
- Peterson,C.L. and Laniel,M.A. (2004) Histones and histone modifications. *Curr. Biol.*, **14**, R546–R551.
- Beato,M. and Eisfeld,K. (1997) Transcription factor access to chromatin. *Nucleic Acids Res.*, **25**, 3559–3563.
- Belikov,S., Gelius,B., Almouzni,G. and Wrangé,O. (2000) Hormone activation induces nucleosome positioning *in vivo*. *EMBO J.*, **19**, 1023–1033.
- Lomvardas,S. and Thanos,D. (2002) Modifying gene expression programs by altering core promoter chromatin architecture. *Cell*, **110**, 261–271.
- Tenen,D.G. (2003) Disruption of differentiation in human cancer: AML shows the way. *Nature Rev. Cancer*, **3**, 89–101.
- Kontaraki,J., Chen,H.H., Riggs,A. and Bonifer,C. (2000) Chromatin fine structure profiles for a developmentally regulated gene: reorganization of the lysozyme locus before trans-activator binding and gene expression. *Genes Dev.*, **14**, 2106–2122.
- Tagoh,H., Himes,R., Clarke,D., Leenen,P.J., Riggs,A.D., Hume,D. and Bonifer,C. (2002) Transcription factor complex formation and chromatin fine structure alterations at the murine *c-fms* (CSF-1 receptor) locus during maturation of myeloid precursor cells. *Genes Dev.*, **16**, 1721–1737.
- Tagoh,H., Schebesta,A., Lefevre,P., Wilson,N., Hume,D., Busslinger,M. and Bonifer,C. (2004) Epigenetic silencing of the *c-fms* locus during B-lymphopoiesis occurs in discrete steps and is reversible. *EMBO J.*, **23**, 4275–4285.
- Mueller,P.R. and Wold,B. (1989) *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science*, **246**, 780–786.
- Pfeifer,G.P., Chen,H.H., Komura,J. and Riggs,A.D. (1999) Chromatin structure analysis by ligation-mediated and terminal transferase-mediated polymerase chain reaction. *Methods Enzymol.*, **304**, 548–571.
- Tormanen,V.T., Swiderski,P.M., Kaplan,B.E., Pfeifer,G.P. and Riggs,A.D. (1992) Extension product capture improves genomic sequencing and DNase I footprinting by ligation-mediated PCR. *Nucleic Acids Res.*, **20**, 5487–5488.
- Chen,H.H., Kontaraki,J., Bonifer,C. and Riggs,A.D. (2001) Terminal transferase-dependent PCR (TDPCR) for *in vivo* UV photofootprinting of vertebrate cells. *Sci. STKE*, **2001**, PL1.
- Dai,S.M., Chen,H.H., Chang,C., Riggs,A.D. and Flanagan,S.D. (2000) Ligation-mediated PCR for quantitative *in vivo* footprinting. *Nat. Biotechnol.*, **18**, 1108–1111.
- Dai,S.M., O'Connor,T.R., Holmquist,G.P., Riggs,A.D. and Flanagan,S.D. (2002) Ligation-mediated PCR: robotic liquid handling for DNA damage and repair. *Biotechniques*, **33**, 1090–1097.
- Follows,G.A., Tagoh,H., Lefevre,P., Hodge,D., Morgan,G.J. and Bonifer,C. (2003) Epigenetic consequences of AML1-ETO action at the human *c-FMS* locus. *EMBO J.*, **22**, 2798–2809.
- Follows,G.A., Tagoh,H., Lefevre,P., Morgan,G.J. and Bonifer,C. (2003) Differential transcription factor occupancy but evolutionarily conserved chromatin features at the human and mouse M-CSF (CSF-1) receptor loci. *Nucleic Acids Res.*, **31**, 5805–5816.
- The Encode consortium (2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, **306**, 636–640.
- Hanlon,S.E. and Lieb,J.D. (2004) Progress and challenges in profiling the dynamics of chromatin and transcription factor binding with DNA microarrays. *Curr. Opin. Genet. Dev.*, **14**, 697–705.