UCLA
UCLA Previously Published Works

Title

Loci specific epigenetic drug sensitivity

Permalink

Journal

ISSN

Authors

Zhang, Thanutra
Pilko, Anna
Wollman, Roy

Publication Date

2020-05-21

DOI

Copyright Information

Peer reviewed

Loci specific epigenetic drug sensitivity

Thanutra Zhang[1], Anna Pilko[1,2] and Roy Wollman [1,2,*]

[1]Institute for Quantitative and Computational Biosciences, University of California, Los Angeles, CA, USA and
[2]Departments of Integrative Biology and Physiology and Chemistry and Biochemistry, University of California UCLA, CA, USA

ABSTRACT

Therapeutic targeting of epigenetic modulators offers a novel approach to the treatment of multiple diseases. The cellular consequences of chemical compounds that target epigenetic regulators (epi-drugs) are complex. Epi-drugs affect global cellular phenotypes and cause local changes to gene expression due to alteration of a gene chromatin environment. Despite increasing use in the clinic, the mechanisms responsible for cellular changes are unclear. Specifically, to what degree the effects are a result of cell-wide changes or disease related locus specific effects is unknown. Here we developed a platform to systematically and simultaneously investigate the sensitivity of epi-drugs at hundreds of genomic locations by combining DNA barcoding, unique split-pool encoding, and single cell expression measurements. Internal controls are used to isolate locus specific effects separately from any global consequences these drugs have. Using this platform we discovered wide-spread loci specific sensitivities to epi-drugs for three distinct epi-drugs that target histone deacetylase, DNA methylation and bromodomain proteins. By leveraging ENCODE data on chromatin modification, we identified features of chromatin environments that are most likely to be affected by epi-drugs. The measurements of loci specific epi-drugs sensitivities will pave the way to the development of targeted therapy for personalized medicine.

INTRODUCTION

The location of a gene on the chromosome is known to affect its expression. Position effect was first observed in *Drosophila* by Muller in 1930 (1,2) and intensively investigated afterward (3–5). Many years after the original work in *Drosophila*, it is now well documented that gene expression levels are influenced by the chromatin environment (2,6–10). Chromatins play a key role in the regulation of gene expression and are responsible for cell maintenance and differentiation. Chromatin regulation is complex and is an area of active research. The three-dimensional structure of the chromatins plays an important role in gene regulation by controlling the accessibility of transcriptional machinery and the spatial proximity of a gene from *cis* regulatory elements such as enhancers. In addition, the specific 3D folding will change the spatial distribution of transcription factors and other regulatory molecules such as lncRNAs. The spatial proximity of these regulatory molecules then plays a key role in controlling gene expression patterns. The three-dimensional structure itself is highly correlated with specific histone and DNA modification patterns. Overall, the complex multi-layered regulation of chromatin on gene expression pattern causes each gene to exist in a unique chromatin environment that plays an important role in determining gene expression distribution, i.e. both its average level as well as population variability (6,9,11).

The proper regulation of gene expression is vital for health and dysregulation of gene expression is associated with a large number of pathologies. Advances in DNA sequencing allow the classification of the specific disease based on the underlying changes of gene expression, the basis of large parts of precision medicine approaches. Given the large knowledge that is accumulating on what changes in gene expression are associated with disease conditions, it is only natural to attempt to correct these pathologies by modification of underlying gene expression patterns. This quest has a long history with initial attempts related to antisense oligos (12). Similarly, the discovery of RNA interference (RNAi), followed by Zn fingers, TALEN, CRISPR (13,14) sparked many attempts to develop therapies with the goal of manipulating gene expression (15). However, despite the conceptual simplicity, translating these concepts into therapy was challenging (16–18).

Given the influence of local chromatin environment on gene expression, strategies that target epigenetic regulators are being investigated. Two main strategies are the pharmacological use of epi-drugs to influence gene expression and targeted approaches for epigenetic editing. Pharmacological approach uses inhibitors to the readers/writers/erasers of epigenetic marks. The pharmacological approach that is

---

*To whom correspondence should be addressed. Tel: +(310)825 4373; Email: rwollman@ucla.edu

being developed to address a wide range of diseases is continuously expanding (19–25). Multiple targeting strategies for epi-drugs are being explored including specific loss and gain of function (26–33), synthetic lethality (34–37), and to overcome drug resistance (37–39). A common theme across these strategies is the use of epi-drugs to manipulate gene expression patterns e.g. suppress oncogenes or activate tumor suppressor genes (40). However, the precision of epi-drugs induced gene expression targeting, i.e. the fraction of overall changes to gene expression that are desired for therapy, is currently very low. This low precision limits the usability of epi-drugs (41–43). The alternative strategy is based on targeted recruitment of epigenetic modulators into specific sites. CRISPR mediated sequence specific targeting of epigenetic regulators is used to cause changes in gene expression pattern of specific loci (44–46). The key advantage of epigenetic engineering is its precision. However, many challenges have to be addressed before these approaches can be translated into the clinic.

Despite the popularity of the use of epi-drugs to cause changes in gene expression patterns, there are many unknowns resulting from gaps in existing measurement capabilities of the effect of epi-drugs on gene expression. Epi-drugs change gene expression due to direct, locus-dependent changes, and indirect or nonspecific effects (40). Existing approaches to identify the direct effects of epi-drugs rely on a combination of RNAseq and multiple ChIPseq to show that gene expression changes are coupled to changes in local histone modification (47). However, the reliance on ChIPseq makes this approach limited as these measurements are only semi-quantitative (48), it is often hard to interpret their functional effects on gene expression (49), and they are challenging to scale to large numbers of samples (50). Therefore, it is currently impossible to rigorously identify changes to gene expression that are due to specific modifications to the chromatin environment and not a result of non-specific or indirect effects. Therefore, there is a gap in current capabilities of mapping the specific and local impact of drug manipulating chromatin modifiers on gene expression.

Here, we developed new measurement technology for the Massive And Parallel Measurement of Epigenetic Drug Sensitivity (MAPMEDS). MAPMEDS is based on a comparison of drug effect at specific locus compared to the drug effect of hundreds of other loci. Statistical comparison of the drug effects on the expression of a reporter fluorescent protein allows the deconvolution of global effects and locus specific sensitivities. MAPMEDS utilizes DNA barcodes to pool together the time consuming step of genomic position identification of the reporters. Split-pool approach enables the mapping of DNA barcodes to individual reporter cell lines. Using MAPMEDS we demonstrate the widespread existence of loci specific sensitivities for three epi-drugs that target histones acetylation, DNA methylation and proteins with bromodomains. By leveraging ENCODE data on the chromatin environment in each location, we show what types of environments are more susceptible to the different epi-drugs. Overall, these results shed light on how epi-drugs cause changes in gene expression, the information that can be used for the development of more precise targeting strategies.

MATERIALS AND METHODS

Cell lines and cell culture

The human K562 cells (Sigma-Aldrich) were grown at 37°C in RPMI 1640 medium (Gibco) supplemented with 10% FBS (Gibco), 1% penicillin–streptomycin (Gibco) and 1% GlutaMAX (100×) (Gibco) under a 95% air and 5% CO_2 atmosphere.

Construction of library reporter plasmid

The base plasmid without barcode was first constructed to contain the following elements. Lentiviral production units include HIV-1 truncated 5′ LTR, HIV-1 packaging signal, HIV-1 Rev response element (RRE), HIV-1 truncated 3′ LTR and central polypurine tract (cPPT). These components allow proper viral packaging and viral integration into host cells. As a transcription unit, we used cytomegalovirus promoter (CMV) to drive expression of the reporter gene encoding yellow-green fluorescent protein (mClover). Woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) is placed after mClover to enhances mRNA stability and protein yield. Ampicillin resistance gene (β-lactamase) is included for selection of plasmid in bacterial cells.

To generate barcoded plasmid libraries, based lentiviral plasmid was cut upstream of the CMV promoter by ClaI restriction enzyme and purified by ethanol precipitation. The inserted cassette of 127-bp-long oligonucleotide containing a random 16-bp-long barcode sequence (repeats of A, T and G), MspI site, primer priming site, and homology arms, were synthesized by Integrated DNA Technology. The assembly reaction of 1:5 vector:insert ratio was carried out for 1 h at 50°C using NEBuilder HIFI DNA assembly kit (New England Biolabs, NEB). Assembly products were electroporated into NEB Turbo Competent *Escherichia coli* (NEB) and then plated on ampicillin-containing medium. Ampicillin resistant colonies were collected and extracted for plasmids using Maxiprep kit (Invitrogen). Ten sampling clones from the agar plate were analyzed by PCR and Sanger sequencing to verify successful cassette insertion and barcode diversity (Primer details in Supplementary Table S1).

Generation of founder cell library and cell lines

Barcoded reporter and third generation lentiviral packaging plasmids were transfected into HEK 293T cells to generate a library of barcoded lentivirus. Viral supernatant was collected and concentrated by Lenti-X-concentrator (Takara) at 48 hour post transfection. K562 cells were transduced with barcoded virus in cultured media supplement with 5 μg/ml polybrene and 20 mM HEPES for 2 h of spinoculation and 24 h of incubation. An m.o.i. of ∼0.01, corresponding to 1% infectivity estimated by flow cytometer, was used to ensure that the majority of cells were labeled with single barcode per cell. Founder cells were selected by fluorescence-activated cell sorting (FACS) at 72 h post transduction. Founder cells were expanded for 2 weeks and split into two pools. In the first pool, cells were subject to mapping the genomic location of barcoded reporter. In

the second pool, cells were sorted into single-cells to establish cell lines of unique barcode.

Identification of genuine barcode list

A library of genuine barcodes in founder cells was first listed. Briefly, barcode region was amplified in first nested PCR from 5 µg of genomic DNA in 50 ul of 20 cycles PCR reaction using Titanium Taq. Barcode amplicons were enriched from genomic DNA using SPRI beads (Beckman Coulter) and further amplified in the second nested PCR for 20 cycles. Illumina adapter was attached to final amplicon, amplified and sequenced on Illumina HiSeq 3000 platform (1 × 50 bp). Sequencing reads were filtered and analyzed using Matlab Bioinformatics Toolbox. To identify genuine barcode, we used the following algorithm. First, we sorted barcodes according to their counts from most frequent to least frequent. Then, mutant versions of each barcode, defined as barcodes within a Hamming distance of 2, were sequentially removed. We consider remaining sequences as 'genuine' barcodes. We recovered 756 genuine barcodes from 3000 sorted founder cells.

Mapping of reporter integration sites

Mapping of reporter integration sites was done by inverse PCR coupled with high-throughput sequencing. Briefly, founder cells were collected and splitted into two replicates. For each replica, 2 µg of genomic DNA was digested with 20 units of MspI (NEB) overnight at 37°C in a volume of 100 µl. Subsequently, three sets of ligation reactions were set up by incubating 600 ng of purified digested DNA with 2 µl of high-concentration T4 DNA ligase (NEB, M0202T) overnight at 4°C in a volume of 400 µl. The ligation reactions were purified by phenol–chloroform isoamyl alcohol extraction and ethanol precipitation. DNA pellets were dissolved in 30 µl of water. Two rounds of PCR were performed to amplify and enrich fragments containing both the barcodes and flanking genomic DNA regions (Primer details in Supplementary Table S1). For the first round of nested PCR, five sets of 25-cycle reaction in a volume of 50 µl were performed using Phusion Hot Start Flex 2× Master Mix (NEB) and 5 µl of ligated products as templates. Amplicon was pooled together, cleaned by DNA Clean & Concentrator kit (Zymo), and diluted in 50 ul of water. For the second round of nested PCR, four sets of 15-cycle reaction in a volume of 50 ul were done with 5 ul of cleaned amplicon from first PCR. Purified sample was further ligated with Illumina adapter, amplified and sequenced on Illumina HiSeq 3000 platform (2 × 150 bp). Sequencing reads were filtered and analyzed using Matlab Bioinformatics Toolbox. The genomic regions associated with genuine barcodes were extracted from mapping reads and aligned against the human genome (hg38) using STAR (51). Detected integration sites from each replicate were compared and assigned to each genuine barcode only if top candidate site from both replicates are identical. Mapping of reporter integration sites were plotted on the ideogram (Figure 2B) using R and karyoploteR package (52). Genome coordinate of reporter integration site was converted to human reference genome (hg19) using UCSC liftOver tool (53) for comparison to ChIP-Seq data.

Combinatorial pool sequencing

Identity of reporter cell lines, linked by DNA barcodes, were simultaneously revealed in a single run using combinatorial pooled sequencing. Clonal numbers were encoded in a form of pooling pattern. To increase decoding accuracy, we designed pooling signature to be unique four selected pools out of total 18 pools. Cells from each clone were splitted into four pools according to the design. Sequentially, genomic DNA from individual pool of mixed clones was extracted and used as templates for PCR to amplify barcode using same procedure described in the method of identification of genuine barcode list. Forward primers of second nested PCR contain 6-bp index DNA to label PCR products from each pool, which allow high-throughput multiplex sequencing (Primer details in Supplementary Table S1). Sequences were filtered and demultiplexed using Matlab Bioinformatics Toolbox. Genuine barcodes from all pools were first listed. For each detected barcode, normalized counts per pools were calculated and pools showing high reads above the threshold were identified. Barcodes with four detected pools were first assigned to the clone showing matched pooling design. Some barcodes were found in more than four pools when sister cells, expanded from one founder cell, were sorted into multiple wells during single-cell sort. A list of merged pooling signature of two unassigned clones was matched with barcodes showing complexed readout. Clones with two inserted barcodes (~2% of the population) were excluded from the library of reporter cell lines.

Epigenetic drug treatment

We first created reference cells expressing both mClover and IRFP670 from multiple integration sites. Briefly, K562 cells were transduced with CMV-IRFP670 lentivirus at high m.o.i. and sorted for IRFP positive cells. Lentiviral transduction of CMV-mClover was followed and dual reporter cells were selected by FACS. Reference cells were cocultured with individual reporter clones. 0.5 million cells of mixed samples were separately treated with 1 µM of JQ1 for 24 h, 250 nM of SAHA for 72 h and 250 nM of 5′ Azacytidine for 72 h. Afterward, cells were collected and measured for expression distribution of mClover and IRFP using BD FACSCelesta flow cytometer. Experiments were done in three replicates per drug treatment per clone. IRFP expression was used to separate control cells from sample cells. To eliminate non loci-specific effects, $\log_{10}$-transformed mclover expression of control cells with epigenetic drug treatment was calibrated to match corresponding expression in DMSO condition. Same adjustment was applied to reporter cells in the same well. the Kolmogorov–Smirnov test was performed by Matlab software to compare histogram similarity of mClover distribution under different conditions.

Validation of combinatorial pool sequencing

For the validation of combinatorial pool sequencing, five clones were randomly chosen and two of them are sister clones, sharing the same barcode. Genomic DNA (200 ng) was used as a template for amplification with a set of validation primers (Primer details in Supplementary Table S1).

PCR products were cleaned by DNA Clean & Concentrator kit (Zymo) and Sanger sequenced to verify the barcode sequences.

RNA extraction and gene expression analysis

K562 cells cultured at 2×10^5 cells/ml were treated with JQ1 at a final concentration of 1 μM for 24 h before the isolation of RNA. Same final concentration of the vehicle, 0.1% DMSO (Sigma), was used to treat the control sample. Total RNA was isolated using RNeasy Mini Kit (Qiagen) with on-column DNase digestion before RNA cleanup. RNA was quantified using Qubit RNA Assay Kit (Thermo Fisher Scientific), and the integrity was assessed on a Tapestation RNA Screentape (Agilent Laboratories). All samples had an RNA integrity number of ≥9. The libraries were sequenced on HiSeq3000 as 50 bp single-end reads. About 40 million reads were obtained from each sample (two control samples and two JQ1 samples). The reads were quality controlled using FastQC and mapped to the human genome using STAR. DESeq2 (54) was used for differential gene expression analysis. Statistical significance of impact of JQ-1 on native genes that are in proximity of loci predicted by our analysis to have loci specific effect was done using a rank permutation test. First, all loci were ranked based on the expected strength of a locus specific effect. Then, the average rank of loci where the nearest gene is significantly impacted by JQ-1 was compared to random permutations of all ranks.

Drug-induced cytotoxicity assessment

K562 cells were stained with Hoechst 33342 for nuclear label and resuspended in imaging media supplemented with propidium iodide for cell death detection. Cells were then seeded in poly-L-lysine coated 96-well plate at a density of 10 000 cells/well. Cell growth and death under different concentrations of drug treatment was monitored by fluorescence imaging every 12 h for 3 days. Imaging analysis used to determine cell growth and death was accomplished using a custom MATLAB code.

Analysis of ChIP-Seq data

Following ChIP-seq data sets were downloaded from ENCODE: H2A.Z (ENCFF191EXE), H3K4me1 (ENCFF526QTS), H3K4me2 (ENCFF118MMT), H3K4me3 (ENCFF715DGL), H3K9ac (ENCFF602QRW), H3K9me1 (ENCFF526UWC), H3K9me3 (ENCFF834YLI), H3K27ac (ENCFF010PHG), H3K27me3 (ENCFF445UCR), H3K36me3 (ENCFF678IWR), H3K79me2 (ENCFF003CLZ) and H4K20me1 (ENCFF143CUR). Fold change over control signals were averaged within a window of 10 kb centered around integration sites using R and Bioconductor packages.

Methylation analysis of CMV promoter

To assess the levels of DNA methylation of CMV promoter, targeted bisulfite conversion was performed. Genomic DNA from clones of interest was extracted and bisulfite converted with Qiagen's EpiTect bisulfite kit according to manufacturer's instructions. Bisulfite converted genomic DNA was used as a template for two rounds of nested PCR using Invitrogen's Phusion U polymerase (Primer details in Supplementary Table S1). Amplicons from all samples were pooled together, purified by DNA Clean & Concentrator kit (Zymo) and deep sequenced. Sequencing reads were filtered and demultiplexed by barcode using Matlab Bioinformatics Toolbox. Changes in cytosine base at CpG sites were converted into a matrix of methylation status.

RESULT

Overview of MAPMEDS

MAPMEDS is based on two key innovations: (i) A split-pool strategy for the creation of cell lines that uses DNA barcodes to identify what DNA labeling each cell line has and what is the genomic integration position of that barcode (Figure 1A–D). Each expression reporter incorporates a unique DNA barcode into a reporter cassette that contains an identical promoter and fluorescent reporter. These barcodes serve as unique identifiers for mapping the genomic locations of reporters and revealing clonal identities without the need for individual genomic extraction. The use of split-pool encoding allows the mapping between isoclonal line expression and the DNA barcode and thereby connecting genomic information with expression measurements. (ii) A new strategy that uses in-well controls and statistical tools enables the separation of gene expression changes into locus-specific and global non-specific (Figure 1E and F). The use of in-well controls minimizes batch effects and is a key to identify locus specific sensitivities. Collectively, these two steps allow the parallel creation of a large number of reporter cell lines and their use to identify loci specific epigenetic drug sensitivities.

MAPMEDS utilizes DNA barcoded expression reporters integrated as a single copy per cell. In order to generate founder cells with one copy of barcoded reporter, lentiviral transduction at low multiplicity of infection (MOI) was used. In short, we first created a library of lentiviral plasmids containing barcode, MspI restriction site, cytomegalovirus (CMV) promoter and GFP variant, mClover (55) (Figure 1A). The barcodes are 16-bp sequences of random adenosine, thymine and guanine. cytosine is excluded to keep barcode intact for future application to examine DNA methylation through bisulfite conversion. Barcoded lentiviruses were packed and transduced at MOI of ∼0.01 into K562 cells, a leukemia cell line with abundant available epigenetic profiles (56) (Figure 1B). Reporter K562 cells were selected at 72 h post-transduction by fluorescent activated cell sorter (FACS) to establish a pool of 3,000 founder cells. Founder cells were expanded for 2 weeks and split into two pools (Figure 1C). One half was used to identify the genomic location using inverse PCR (6). The other half was used to establish individual isoclonal lines through single cell sorting. Once the cell lines were established, we utilized a combinatorial pooling approach to combine all the cell lines into a small number of pooled samples that were used to identify the barcode identity, and hence the genomic integration site, of the ORFs for all cell lines (Figure 1D). The result of this procedure is a library of clonal cell lines that
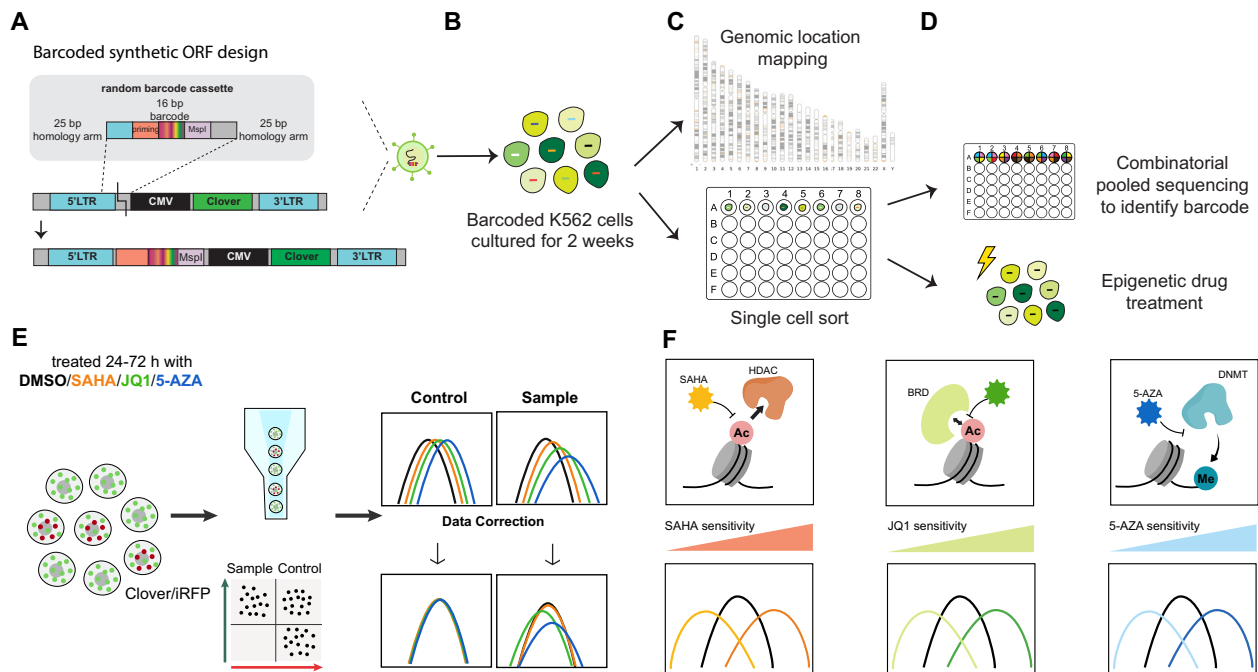
Figure 1. An overview of MAPMEDS. (**A**) Schematic structure of barcoded lentiviral constructs. The library vector contains short barcode, CMV promoter and mClover fluorescent protein as a reporter. The barcode is random 16-bp-long DNA with repeats of A, G and T. MspI restriction site is integrated upstream CMV promoter for genomic location mapping. (**B**) Barcoded lentivirus was packed and transduced into K562 cells at low MOI to create founder cells with singly integrated reporter. (**C, D**) Barcoded founder cells, selected by flow cytometry, were expanded for two weeks and split into two pools. Cells in the first pool were collected for locating reporter integration site. Founder cells in the second pool were sorted into 96-well plates to establish clonal cell lines. Barcode of each clone was simultaneously identified by split-pool encoding and deep sequencing. The library of characterized reporter clones is a useful resource to examine loci specific epigenetic drug sensitivity. (**E**) Loci specific effects were decoupled from global effects through mixing individual barcoded clones of interests with reference cells expressing mClover and IRFP from multiple integration sites. Co-cultured cells were treated with SAHA, JQ1 and 5-AZA for 24–72 h. Expression of reporter proteins were measured by flow cytometers. Distribution of mClover expression in reference cells was used to remove global effects of drugs. (**F**) A cartoon illustrating known mechanisms of actions of SAHA, JQ1 and 5-azacytidine.

can be used as a powerful resource to examine drug sensitivity at diverse epigenetic environments.

In order to isolate locus-specific changes from any global changes in gene expression patterns, we implemented a new strategy using in-well controls and statistical tools. A reference 'non-specific' cell population was created by using a polyclonal population of multiple integration cells so that the overall population has thousands of integration sites of same reporter cassettes. Far-red fluorescent protein iRFP670 (57) was used to mark this reference population. We split these control cells and co-cultured with each target cell line in multiwell plates.

Integration landscapes of reporters

To map the integration sites of reporters, we split half of founder cells into three sub-pools and further expanded the population. The first pool was used to reveal a list of genuine barcodes. We detected 756 candidate genuine barcodes after two weeks of culture. Two other sub-pools are technical replicates for locating reporter integration sites by an inverse PCR method coupled to paired-end high-throughput sequencing (Figure 2A). In short, genomic DNA from each pool was isolated, digested with MspI enzyme and self-ligated. Barcodes and adjacent genomic DNA were amplified and deep sequenced. After barcode demultiplexing, genomic sequence was mapped to human genome assembly GRCh38. Genomic coordinate was assigned to corresponding barcode when mapped results from two replicates are matched. We observed reporter integration throughout the genome (Figure 2B, Supplementary Figure S1A) with enriched pattern similar to previous study (Supplementary Figure S1B) (8). Reporters were integrated at various genomic environments serving as diversified resources to study epigenetic drug sensitivity.

Scalability and robustness of MAPMEDS

Pooled-sample sequencing is a cost-effective and practical strategy for many studies, especially the ones related to the discovery of rare mutation and single nucleotide polymorphism associated diseases (58–60). Combinatorial pooled sequencing is an extension of standard pooled sampling where each sample exists in a few pools creating a many to many mapping between experimental conditions and samples. Combinatorial pooling improves the sensitivity and robustness of pooled sequencing since it includes built in error corrections. It reduces the cost and time for library preparation exponentially (61–63).

Conceptually, in combinatorial pooled sequencing, the identity of each sample is encoded in the composition of pools and this pooling pattern serves as a reference for decoding sequences belonging to corresponding sample (Figure 3A). Clonal lines are mixed into few pools according to
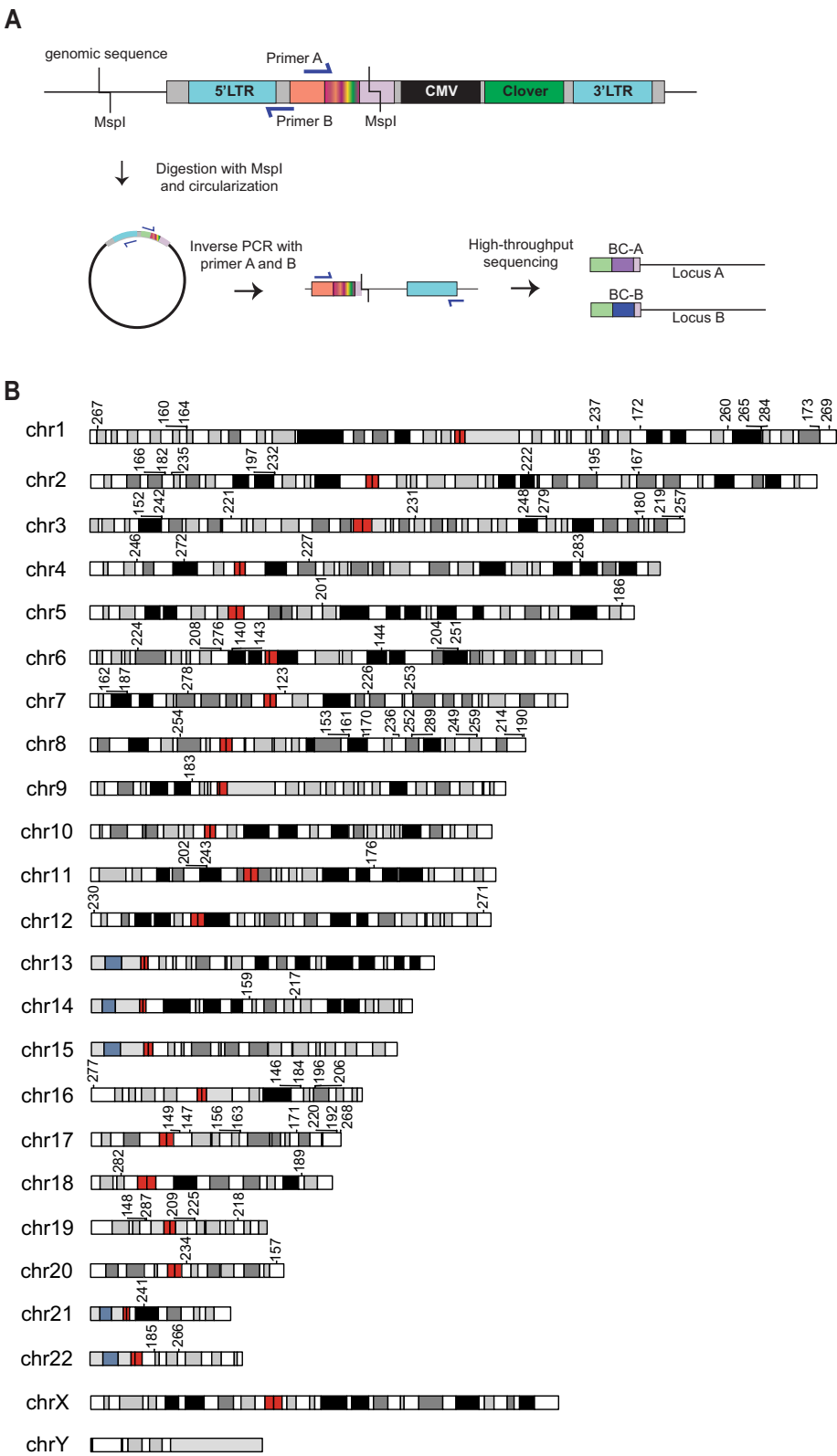
Figure 2. Diverse insertion landscapes of barcoded reporter. (**A**) Reporter mapping by inverse PCR. Genomic DNA of founder cells was extracted, digested with restriction enzyme MspI and self-ligated to stitch barcode with its neighboring genome. Ligated product was amplified and followed by next generation sequencing. (**B**) Ideogram plot displaying reporter integration sites of individual clones in the library. Centromere position is indicated in red and stalk is marked in light blue. Heterochromatic region, which tends to be rich with adenine and thymine and relatively gene-poor, is represented by black and variation of grey. R-band in white on the ideogram is less condensed chromatin that is transcriptionally more active.
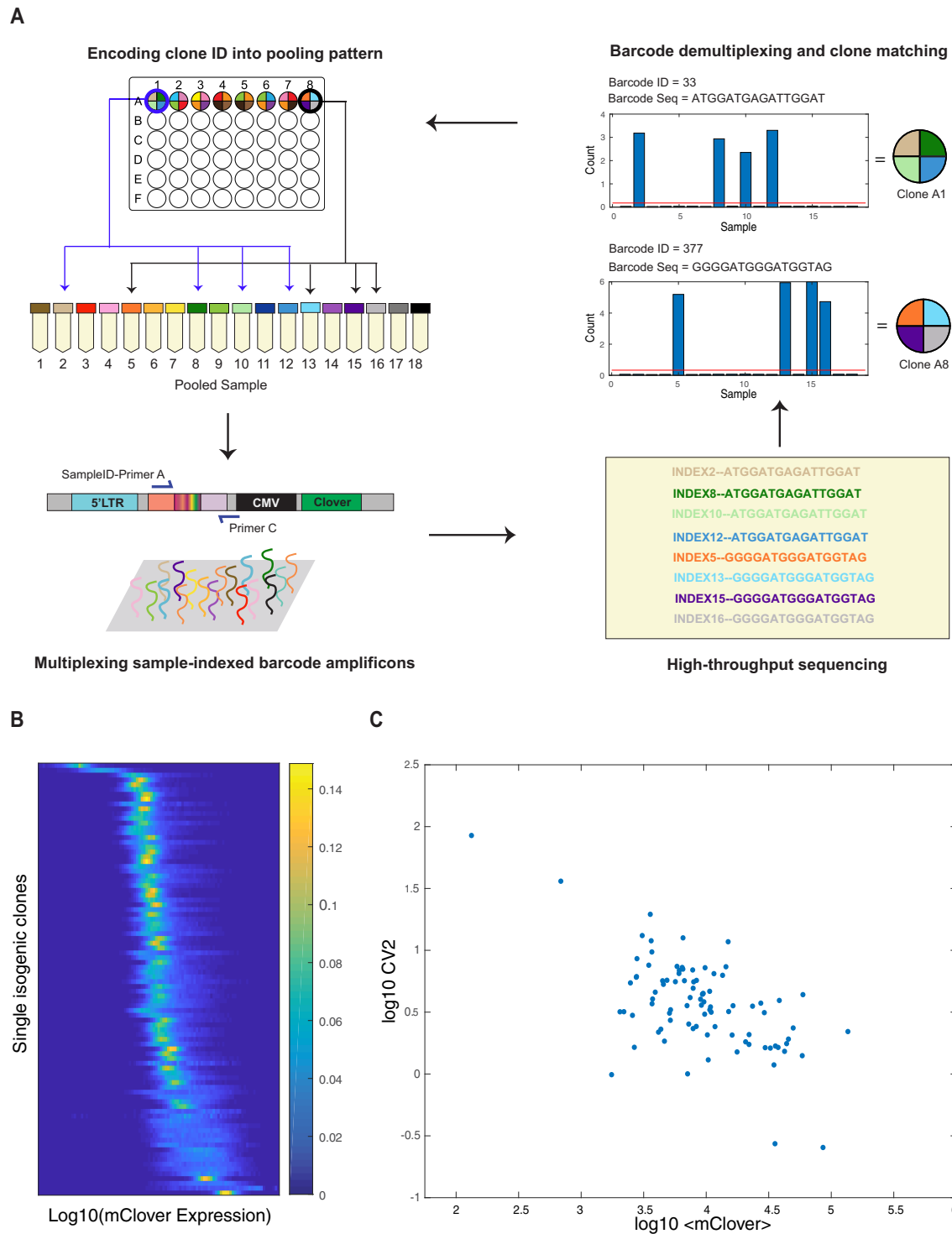
Figure 3. Combinatorial pooling massively and parallely identify barcodes of individual clones. (**A**) The combinatorial pooled sequencing involves encoding and decoding steps. Each individual clone was split into four out of 18 pooled samples according to the designs. Genomic DNA from each pooled sample was extracted and barcodes were amplified and labeled with 6-bp sample index. Amplicons from all samples were mixed together and prepared for NGS. Detected barcodes were deconvoluted to match barcode ID with clone ID. (**B**) The expression distribution of mClover protein was measured by high-throughput flow cytometer and displayed by stacked probability density function. Each row in the heatmap represents a single histogram from a single clonal line with the probability density function colorcoded. (**C**) A scatter plot of reporter expression noise, measured as log-transformed squared coefficient of variation (CV^2), and mClover mean across all examined positions affirms unique chromatin environments across the genome.

a predefined design. Each individual pool is genomic DNA extracted, barcode amplified and index tagged using nested PCR. Amplicons from all pools are mixed together and sequenced. Subsequently, sequencing data is sorted by barcodes and their appearance patterns determined by the well indexes. The measured patterns are compared with the design pooling signatures to match barcodes with clone numbers.

The use of combinatorial pooling circumvents the need for individual genomic extraction and PCR per clones which significantly reduces the cost of reagents and hands-on time for sequencing preparation. In our design, each clone is mapped into 4 selected pools from total of 18 pools and this set-up allows up to 18-choose-4 or 3060 sample identification. This approach is highly scalable because sample size can be easily increased by adjusting number of pool and bit of code. We checked the accuracy of barcode identification using combinatorial pool sequencing by targeted PCR and Sanger sequencing. For five randomly chosen clones, the barcode deconvolution was all corrected which confirm the robustness of our method.

Chromosomal position effects on protein expression distribution

To support the notion that each integration site exists in a different chromatin environment, we examined the positional effect of our expression reporters. Across nearly a hundred positions, we observed variable expression distribution of mClover protein (Figure 3B). Some clones are completely silent while many are highly expressed with expression average of ∼1000-fold higher than the lowest-expressing cells (Figure 3C). This variation range is comparable to a study that measured averaged mRNA expression across mouse genome (6), but higher than other studies carried out in bacteria or yeast (7,9–10). Broader magnitude of positional effect in expression level detected in mammalian genomes may come from their large and complex genome organization and the discrepancy in experimental design and techniques.

Our observation on variable expression distribution confirms that the location where reporter gene inserted affects its expression and differences in genomic landscapes and epigenetic profiles are suggested to explain such differential expression in several studies. For example, lamina-associated domain and chromatin compaction significantly attenuate transcriptional activity (6) and certain histone marks, including H3K36me3, are correlated with expression level of the reporters (9,10). Intuitively, epigenetic drugs that target chromatin regulator should also be impacted by distinct genomic environments. However, such loci specific sensitivity has not been previously measured and this motivates us to systematically measure positional effects on the sensitivity of epigenetic drugs using our library of isogenic clones established and characterized by MAPMED.

Epigenetic drugs show position-dependent sensitivity

As proof of concept, we chose three epigenetic drugs representing three mechanisms of inhibition. Suberoylanilide hydroxamic acid (SAHA) is histone deacetylases (HDACs) inhibitor (64,65) and effective in the treatment of several types of cancer (66–68). JQ1 is a small-molecule inhibitor of BRD2 and BRD4, members of the bromodomain and extra-terminal domain (BET) protein family. JQ1 competitively blocks the binding of bromodomain proteins and acetylated chromatin which results in transcriptional attenuation (69,70). JQ1 has been reported as a promising cancer therapeutic strategy in several cancers (71–73). Azacytidine is a cytidine analogue that inhibits DNA methylation through loss of DNA methyltransferase (DNMT) activity (74). Azacitidine was approved by the U.S. Food and Drug Administration (FDA) for the treatment of all subtypes of myelodysplastic syndrome (MDS) since 2004 (75). In advance of epi-drug effect measurement, the cytotoxicity test was first performed to select appropriate concentration of drug that is high enough to slow K562 proliferation by ∼50% but limit cell death at 60 h to be less than 10% (Supplementary Figure S2). Three drugs at selected concentration were added to co-cultured cells and incubated for 24–72 h before sample collection and fluorescent quantification.

The expression level of mClover and iRFP was measured using high throughput sampler (HTS) flow cytometer. Cells from each sample were separated into reference and target cells based on iRFP gating. To quantify the site specific effect, we normalized the expression levels of the sample cells by the average change between the drug and DMSO in the control population. After correction, the site specific effects of a drug that are not captured by reference population were quantified by the Kolmogorov–Smirnov (KS) statistics that compares distribution similarity between sample cells under drug and sample cells with DMSO (Figure 4A–C). Any effects that are only site specific are effectively measured as changes beyond those also occur in the reference polyclonal population. Differential magnitudes of drug sensitivity were observed across examined locations. Most positions are either hyposensitive or hypersensitive to only specific epigenetic drugs in a non-uniform pattern across the genome (Figure 4D–F) which further suggests the unique chromatin environment at different genomic loci.

The distributions of KS statistics in control cells fall within the three standard deviation limits (Supplementary Figure S3). Therefore, we used this criteria to identify the number of 'hits' per drug as it provides a non-parametric estimate of our false discovery rate to be <0.015. Our drug screening shows ∼44.3% and ∼29.5% of the positions that have down-regulated and up-regulated reporter expression after SAHA treatment, ∼64% and ∼24.5% in JQ1 and 34.4% and 52.5% in 5-azacytidine respectively. Overall, our data demonstrates a chromosomal position effects on epigenetic drug sensitivity.

Analysis of histone modification profiles identifies chromosomal environments susceptible to JQ1 sensitivity

We compared our maps of epigenetic drug sensitivity to a collection of available epigenomic maps from K562 cells focusing on histone modifications. ChIP-seq signals, expressed by fold change over control, in a window of 10 kb around barcodes of interest were considered. We note that histone modification profiles were mapped in K562
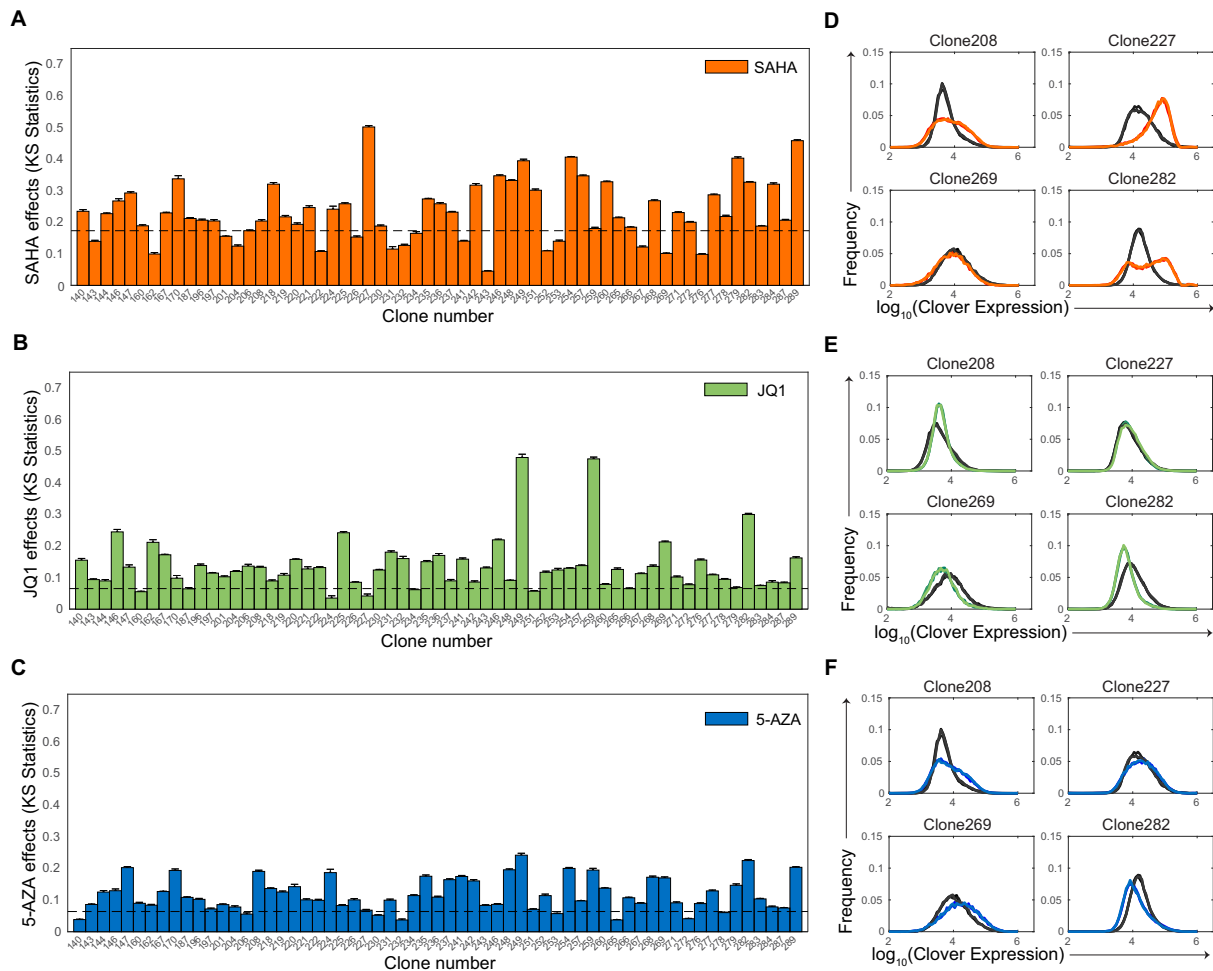
Figure 4. Chromosomal position effects influence magnitudes of epigenetic drug effects. (**A–C**) A bar graph of loci-specific drug effects scored by the two-sample Kolmogorov–Smirnov (KS) test between DMSO and epigenetic drug treatment of SAHA (A), JQ1 (B) and 5-azacytidine (C). The dashed line represents the three standard deviations of KS statistics in control cells. (**D–F**) Histogram plot showing mClover distributions of selected clones after the treatment of DMSO (black), SAHA (orange), JQ1 (green) and 5-azacytidine (blue). Each position shows unique responses to different drugs.

cells without any genetic engineering. Insertion of reporters potentially change the pre-existing epigenetic landscape or cause sequence-specific or protein-specific interactions between regional chromatin and synthetic ORF. However, previous studies suggest that integrated reporters generally do not perturb the chromatin landscape but adopt the local chromatin state (10,76). Complex sequence- or protein-specific interactions between local chromatin and synthetic reporters were not observed in previous study as well (11).

After selecting top JQ1-sensitive clones with distinct downregulation and upregulation of reporter expression (Figure 5A), we compared their epigenetic profiles. Interestingly, we found JQ1 sensitivity is associated with certain histone modifications (Figure 5B). Structural study reveals the binding of JQ1 to the acetyl-lysine binding pocket of BET bromodomains (69). Such competitive binding disrupts bromodomain/acetyl histone interaction and therefore transcriptional activation. Genomic regions enriched in acetylated histones were hypothesized to display higher magnitude of reporter downregulation. Indeed, we observed that clones with lower expression after JQ1 treat-

ment show higher enrichment of H3K9ac and H3K27ac. Moreover, we also found that these clones have significantly higher enrichment of H2A.Z. This data supports previous studies demonstrating that BRD2 interacts with H2A.Z to mediate transcription initiation (77,78) and thus suggests that our assay recapitulated well-established results.

If indeed the changes in expression patterns reflect changes to underlying chromatin environment we expect to see that native genes in proximity to these sensitive loci will also show differential gene expression. The regulation of native gene expression patterns is highly complex with multiple layers of feedback regulation. Therefore, we do not expect all genes to directly change in the same way that a synthetic reporter does. Nonetheless, the loci specific sensitivities should reflect changes to native genes. To test that we performed a standard differential expression change using RNA-seq after treatment with JQ1 (Figure 5C). For the 61 sites of integration of the synthetic reporters we checked whether the closest gene showed differential response to JQ1. In the majority of these cases, the integration site was in the intronic region of the gene. We found statistically sig-
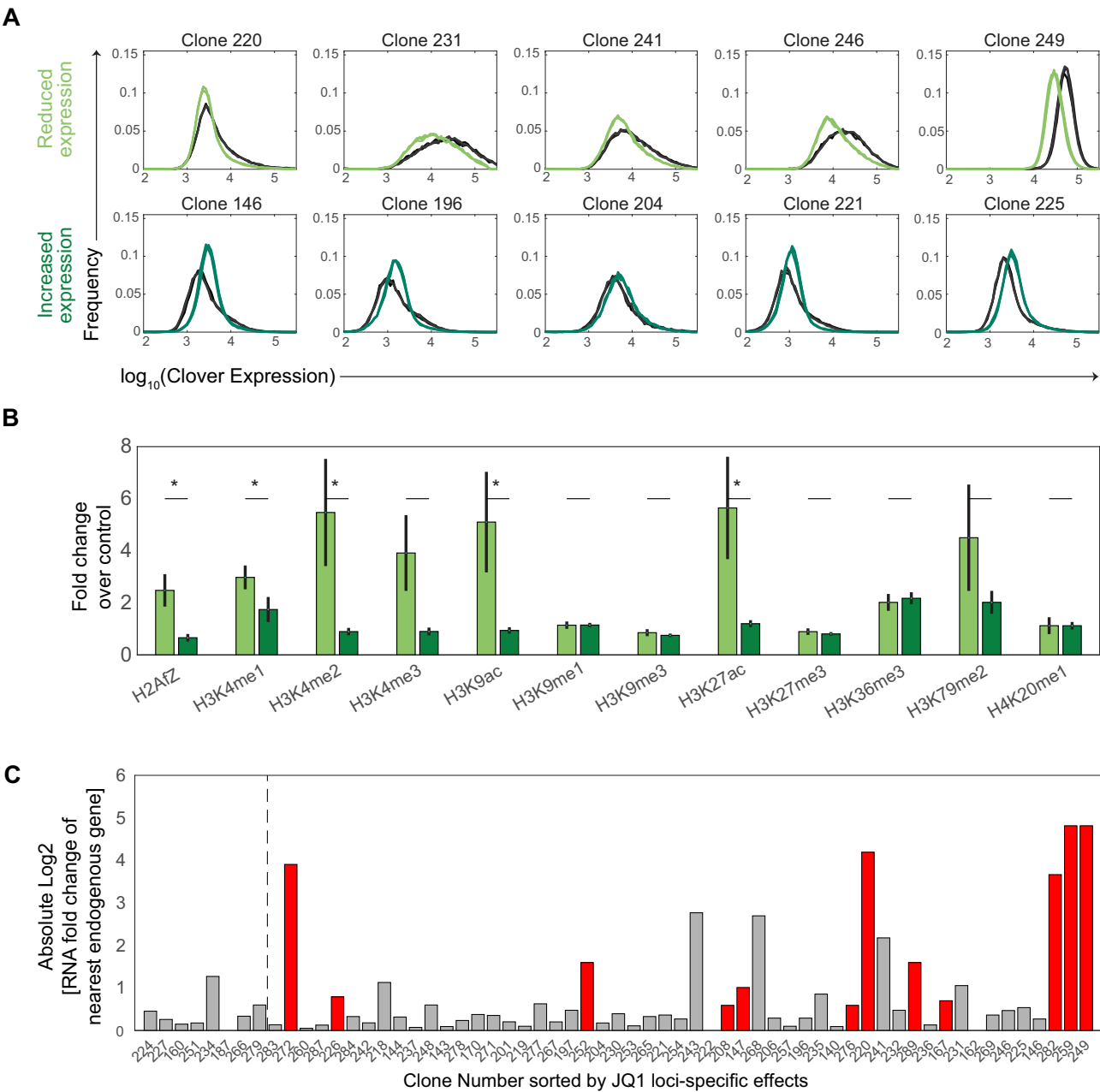
Figure 5. Chromatin environment influences sensitivity of bromodomain inhibitor to expression alteration. (**A**) Examples of mClover distribution from clones showing significant reduced (top row) and increased (bottom row) mClover expression to JQ1 drug. (**B**) Bar graph shows histone enrichments based on data from the ENCODE consortium for comparison of positions displaying reporter down-regulation (light green) versus those exhibiting up-regulation (dark green) after JQ1 treatment. Fold change over control of each histone mark was averaged within a window of 10 kb. The *P*-values were determined by two-sample *t*-test. (**C**) Bar graph of the log2 fold change of the expression of the nearest native gene to each of the synthetic reporter integration sites. Clones are sorted based on the magnitude of the effect at each locus. Red bars indicate a statistically significant change in gene expression determined using DEseq2.

nificant enrichment of native gene JQ1 sensitivity for sites we identified as JQ1 sensitive using the synthetic reporter assay (*P*-value < 0.001 rank permutation test).

Differential sensitivity to 5-AZA treatment happened through DNA methylation-independent mechanism

Considering azacytidine as a well-known inhibitor of DNA methylation, it is likely that genomic regions with hyper-

methylated promoter will respond to such epigenetic drugs the most. Moreover, CMV promoter used in our reporter is highly enriched in CpG sites which should be susceptible to DNA methylation. Therefore, we hypothesized that reporters inserted at diverse genomic environments will have different DNA methylation levels and result in differential drug sensitivity. To test this hypothesis, we examined DNA methylation at CMV promoter, which contains 30 CG sites, using target bisulfite sequencing (Figure 6A).
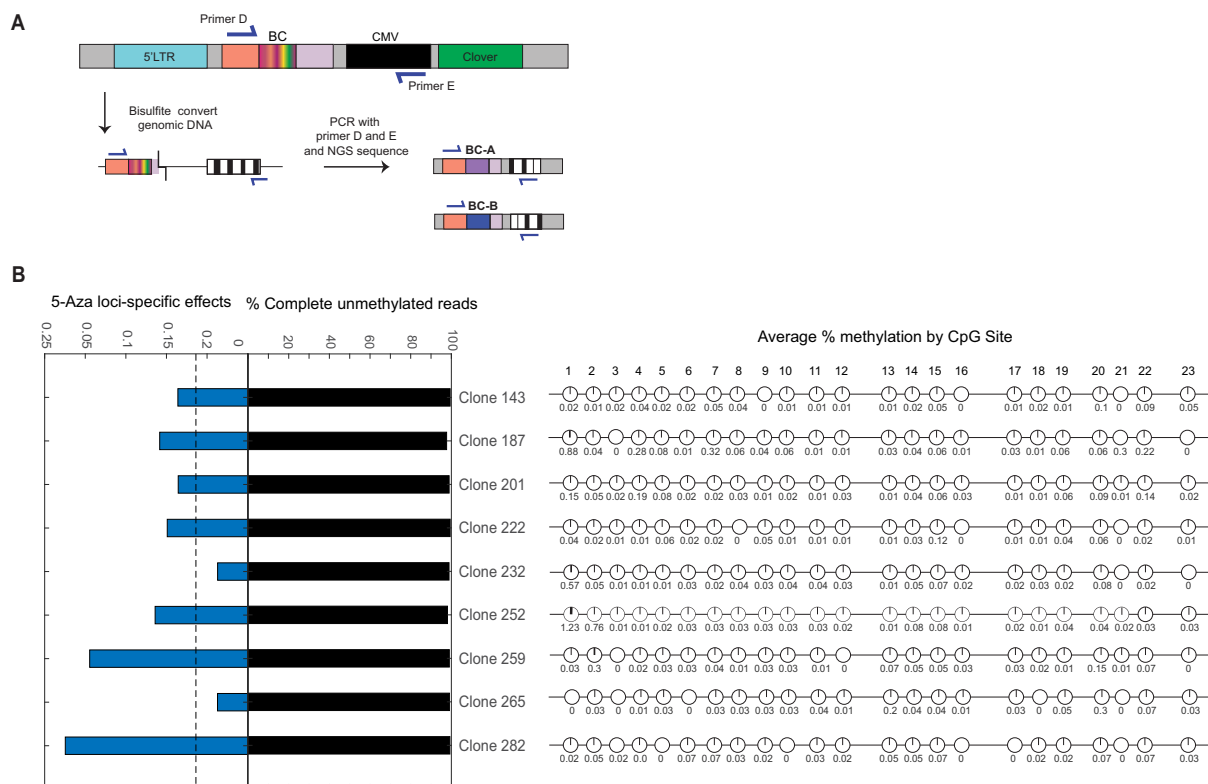
Figure 6. Chromosomal position effects 5-azacytidine sensitivity through DNA methylation-independent mechanism. (**A**) DNA methylation profiles of selected clones were simultaneously revealed by bisulfite conversion, targeted PCR and deep sequencing. (**B**) Loci-specific drug effects were compared with DNA methylation levels on CMV promoter revealed by bisulfite sequencing. About ten thousand of bisulfite sequencing reads were archived from representative clones exhibiting differential response to 5-azacytidine. The dashed line represents the threshold of significant responses.

We did not identify changes in the methylation status of the reporter even in loci where we observed loci specific changes in expression (Figure 6B). The CMV promoters are mostly unmethylated and indistinguishable (Figure 6B). These results are not surprising as by design we could only focus on CMV promoters that are expressed and therefore are not methylated in their promoters. Our data suggests that differential sensitivity to 5-azacytidine treatment may not result from the direct inhibition of promoter methylation but instead reflect changes in cell regulatory machinery that has stronger effects on these loci than average. This observation is in agreement with previous studies that report 50–60% of induced genes by 5-Aza-CdR did not have CpG islands within their 5′region (79,80).

DISCUSSION

Here, we developed a new method MAPMEDS that can identify loci specific drug sensitivity in a robust and scalable manner. The method uses DNA barcoding to generate cell lines with known integration position of an expression reporter at genome scale and includes a statistical procedure to quantify the locus specific effect that a drug has on gene expression. We used MAPMEDS to evaluate the position specific effects of three common epi-drugs. We found that up to 80% of positions change in a manner that is different than average including both hyper and hypo sensi-

tivity. Through analysis of the chromatin features that are enriched in sites with hyper/hypo drug sensitivities we were able to characterize what aspects of chromatin environment make it more (or less) sensitive to a specific drug. The development of MAPMEDS have both translational and basic science implications.

Locus-specific sensitivity measurements will support the development of new treatment strategies that use existing drugs and the development of new, more precise drugs. Identification of correct dosage of epi-drugs is challenging (81–83). Comparison of the dose-response curve of the overall change in gene expression to the dose-response changes in gene expression that are due to locus-specific modification will help identifying drug concentrations that maximally impact target genes while limiting non-specific effects. Similarly, it will be possible to improve the precision of drug combinations (21,84). New drugs and lead compounds could be identified based on predefined desired locus-specific changes in expression patterns. For example, in breast cancer that develops resistance to PI3K therapy, it was shown that co-treatment with a bromodomain inhibitor JQ1 helps mitigate drug resistance since it silences the compensatory upregulation of RTKs (85). The use of JQ1 to achieve such desired effect is limited by the fact that JQ1 has very broad effects on expression changes across the genome. The tools we propose to develop will allow screening for new compounds and JQ1 derivatives that maximizes

the effects on the desired genes such as EGFR and INSR while limiting other undesirable changes.

Precision medicine is based on stratification of patients and assigning specific therapies based on the molecular information often associated with the key aberrant pathways. Therefore, in many cases, the nature of needed changes in gene expression are known. Treatment is limited due to the lack of therapies that can cause such desired changes in gene expression. Current manipulations of gene expression patterns using epigenetic targeting drugs such as JQ1 are imprecise (41,42). The new measurement technology developed here will support the future development of more precise epigenetic drug-mediated gene expression targeting. The reduction in side effects and the ability to screen for locus-specific changes in gene expression will lead to a large array of new therapies.

From basic science perspective, our understanding of chromatin regulation of gene expression is far from complete. MAPMEDs has the capacity to generate large datasets that look at changes across many positions and epigenetic drugs. Such large dataset will provide key insights into how changes in the local chromatin environment can affect gene expression in a manner isolated from any global changes. The use of DNA barcodes as part of MAPMEDS enables the pooled measurement of changes to local chromatin environment at scale. These data will provide invaluable insight into chromatin regulation.

MAPMEDS have two unique aspects that set it apart from other measurement approaches such as TRIP and BHIV (6,8) that aimed at measuring positional effects at scale. Unlike other approaches, MAPMEDS is based on the library creation of cell lines. The use of library of cell lines provides single cell data on changes in expression variability. Indeed, many of the locus specific drug effects we saw did not simply shifted the population but changes the shape of the distribution. These effects would have been missed with bulk population measurements. Additionally, once the initial work in creating the cell line library was invested, the measurement of the drug effects are straightforward and can be scaled to large drug libraries. The work presented here was based on one such library of cell lines and future work is needed to explore larger numbers of loci, cell types, and promoters. Other approaches such as TRIP and BHIV will require full RNA sequencing for each drug tested. The down side of the cell line library approach is that it is hard to scale it to more than a few hundred sites. The small scale limits the overall information and could generate a non-uniform representation of loci in the genome, especially if generating these reporter cell lines is more difficult in some genomic regions. However, recent advances in in-cell barcodes make it possible to generate the cell line library in pooled format, opening the way to a few orders of magnitude increase in the number of positions that can be measured. Future development of MAPMEDS to include these in-cell barcodes will further increase it's utility as a platform for the discovery of more precise and useful epi-drugs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Muller,H.J. (1930) Types of visible variations induced by X-rays inDrosophila. *J. Genet.*, **22**, 299–334.
2. Henikoff,S. (1990) Position-effect variegation after 60 years. *Trends Genet.*, **6**, 422–426.
3. Weiler,K.S. and Wakimoto,B.T. (1995) Heterochromatin and gene expression in Drosophila. *Annu. Rev. Genet.*, **29**, 577–605.
4. Ottaviani,A., Gilson,E. and Magdinier,F. (2008) Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie*, **90**, 93–107.
5. Elgin,S.C.R. and Reuter,G. (2013) Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harb. Perspect. Biol.*, **5**, a017780.
6. Akhtar,W., de Jong,J., Pindyurin,A.V., Pagie,L., Meuleman,W., de Ridder,J., Berns,A., Wessels,L.F.A., van Lohuizen,M. and van Steensel,B. (2013) Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell*, **154**, 914–927.
7. Bryant,J.A., Sellars,L.E., Busby,S.J.W. and Lee,D.J. (2014) Chromosome position effects on gene expression in Escherichia coli K-12. *Nucleic Acids Res.*, **42**, 11383–11392.
8. Chen,H.-.C., Martinez,J.P., Zorita,E., Meyerhans,A. and Filion,G.J. (2017) Position effects influence HIV latency reversal. *Nat. Struct. Mol. Biol.*, **24**, 47–54.
9. Chen,X. and Zhang,J. (2016) The genomic landscape of position effects on protein expression level and noise in yeast. *Cell Syst.*, **2**, 347–354.
10. Chen,M., Licon,K., Otsuka,R., Pillus,L. and Ideker,T. (2013) Decoupling epigenetic and genetic effects through systematic analysis of gene position. *Cell Rep.*, **3**, 128–137.
11. Maricque,B.B., Chaudhari,H.G. and Cohen,B.A. (2018) A massively parallel reporter assay dissects the influence of chromatin structure on cis-regulatory activity. *Nat. Biotechnol*, doi:10.1038/nbt.4285.
12. Stephenson,M.L. and Zamecnik,P.C. (1978) Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 285–288.
13. Shankar,S., Sreekumar,A., Prasad,D., Das,A.V. and Pillai,M.R. (2018) Genome editing of oncogenes with ZFNs and TALENs: caveats in nuclease design. *Cancer Cell Int.*, **18**, 169.
14. Ho,B.X., Loh,S.J.H., Chan,W.K. and Soh,B.S. (2018) In vivo genome editing as a therapeutic approach. *Int. J. Mol. Sci.*, **19**, doi:10.3390/ijms19092721
15. Wood,M., Yin,H. and McClorey,G. (2007) Modulating the expression of disease genes with RNA-based therapy. *PLoS Genet.*, **3**, e109.
16. Paroo,Z. and Corey,D.R. (2004) Challenges for RNAi in vivo. *Trends Biotechnol.*, **22**, 390–394.

17. Dutta,T. (2010) Challenges in siRNA/ shRNA delivery and development of RNAi therapeutics for cancer. *J. Bioequiv. Bioavail.*, **6**, 1130–1146.

18. Boudreau,R.L., Rodríguez-Lebrón,E. and Davidson,B.L. (2011) RNAi medicine for the brain: progresses and challenges. *Hum. Mol. Genet.*, **20**, R21–R27.

19. Schnekenburger,M., Florean,C., Dicato,M. and Diederich,M. (2016) Epigenetic alterations as a universal feature of cancer hallmarks and a promising target for personalized treatments. *Curr. Top. Med. Chem.*, **16**, 745–776.

20. Pfister,S.X. and Ashworth,A. (2017) Marked for death: targeting epigenetic changes in cancer. *Nat. Rev. Drug Discov.*, **16**, 241–263.

21. Raynal,N.J.-.M., Da Costa,E.M., Lee,J.T., Gharibyan,V., Ahmed,S., Zhang,H., Sato,T., Malouf,G.G. and Issa,J.-P.J. (2017) Repositioning FDA-Approved drugs in combination with epigenetic drugs to reprogram colon cancer epigenome. *Mol. Cancer Ther.*, **16**, 397–407.

22. Jones,P.A., Issa,J.-.P.J. and Baylin,S. (2016) Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.*, **17**, 630–641.

23. Bertino,E.M. and Otterson,G.A. (2011) Romidepsin: a novel histone deacetylase inhibitor for cancer. *Expert Opin. Investig. Drugs*, **20**, 1151–1158.

24. Lee,H.-.Z., Kwitkowski,V.E., Del Valle,P.L., Ricci,M.S., Saber,H., Habtemariam,B.A., Bullock,J., Bloomquist,E., Shen,Y.L., Chen,X.-H. *et al.* (2015) FDA Approval: Belinostat for the treatment of patients with relapsed or refractory peripheral T-cell lymphoma. *Clin. Cancer Res.*, **21**, 2666–2670.

25. Mann,B.S., Johnson,J.R., Cohen,M.H., Justice,R. and Pazdur,R. (2007) FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist*, **12**, 1247–1252.

26. Raynal,N.J.-.M., Lee,J.T., Wang,Y., Beaudry,A., Madireddi,P., Garriga,J., Malouf,G.G., Dumont,S., Dettman,E.J., Gharibyan,V. *et al.* (2016) Targeting calcium signaling induces epigenetic reactivation of tumor suppressor genes in cancer. *Cancer Res.*, **76**, 1494–1505.

27. Rathert,P., Roth,M., Neumann,T., Muerdter,F., Roe,J.-.S., Muhar,M., Deswal,S., Cerny-Reiterer,S., Peter,B., Jude,J. *et al.* (2015) Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature*, **525**, 543–547.

28. Zhu,H., Bengsch,F., Svoronos,N., Rutkowski,M.R., Bitler,B.G., Allegrezza,M.J., Yokoyama,Y., Kossenkov,A.V., Bradner,J.E., Conejo-Garcia,J.R. *et al.* (2016) BET bromodomain inhibition promotes anti-tumor immunity by suppressing PD-L1 expression. *Cell Rep.*, **16**, 2829–2837.

29. Chiappinelli,K.B., Strissel,P.L., Desrichard,A., Li,H., Henke,C., Akman,B., Hein,A., Rote,N.S., Cope,L.M., Snyder,A. *et al.* (2016) Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell*, **164**, 1073.

30. Karpf,A.R., Peterson,P.W., Rawlins,J.T., Dalley,B.K., Yang,Q., Albertsen,H. and Jones,D.A. (1999) Inhibition of DNA methyltransferase stimulates the expression of signal transducer and activator of transcription 1, 2, and 3 genes in colon tumor cells. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 14007–14012.

31. Puissant,A., Frumm,S.M., Alexe,G., Bassil,C.F., Qi,J., Chanthery,Y.H., Nekritz,E.A., Zeid,R., Gustafson,W.C., Greninger,P. *et al.* (2013) Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.*, **3**, 308–323.

32. Rocchi,P., Tonelli,R., Camerin,C., Purgato,S., Fronza,R., Bianucci,F., Guerra,F., Pession,A. and Ferreri,A.M. (2005) p21Waf1/Cip1 is a common target induced by short-chain fatty acid HDAC inhibitors (valproic acid, tributyrin and sodium butyrate) in neuroblastoma cells. *Oncol. Rep.*, **13**, 1139–1144.

33. Brueckner,B., Garcia Boy,R., Siedlecki,P., Musch,T., Kliem,H.C., Zielenkiewicz,P., Suhai,S., Wiessler,M. and Lyko,F. (2005) Epigenetic reactivation of tumor suppressor genes by a novel small-molecule inhibitor of human DNA methyltransferases. *Cancer Res.*, **65**, 6305–6311.

34. Kadoch,C. (2016) Lifting Up the HAT: synthetic lethal screening reveals a novel vulnerability at the CBP–p300 axis. *Cancer Discov.*, **6**, 350–352.

35. Pfister,S.X., Markkanen,E., Jiang,Y., Sarkar,S., Woodcock,M., Orlando,G., Mavrommati,I., Pai,C.-C., Zalmas,L.-P., Drobnitzky,N. *et al.* (2015) Inhibiting WEE1 selectively kills histone H3K36me3-Deficient cancers by dNTP starvation. *Cancer Cell*, **28**, 557–568.

36. Oike,T., Ogiwara,H., Tominaga,Y., Ito,K., Ando,O., Tsuta,K., Mizukami,T., Shimada,Y., Isomura,H., Komachi,M. *et al.* (2013) A synthetic Lethality–Based strategy to treat cancers harboring a genetic deficiency in the chromatin remodeling factor BRG1. *Cancer Res.*, **73**, 5508–5518.

37. Bitler,B.G., Aird,K.M., Garipov,A., Li,H., Amatangelo,M., Kossenkov,A.V., Schultz,D.C., Liu,Q., Shih,I.-M., Conejo-Garcia,J.R. *et al.* (2015) Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat. Med.*, **21**, 231–238.

38. Sharma,S.V., Lee,D.Y., Li,B., Quinlan,M.P., Takahashi,F., Maheswaran,S., McDermott,U., Azizian,N., Zou,L., Fischbach,M.A. *et al.* (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, **141**, 69–80.

39. Vinogradova,M., Gehling,V.S., Gustafson,A., Arora,S., Tindell,C.A., Wilson,C., Williamson,K.E., Guler,G.D., Gangurde,P., Manieri,W. *et al.* (2016) An inhibitor of KDM5 demethylases reduces survival of drug-tolerant cancer cells. *Nat. Chem. Biol.*, **12**, 531–538.

40. de Groote,M.L., Verschure,P.J. and Rots,M.G. (2012) Epigenetic Editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. *Nucleic Acids Res.*, **40**, 10596–10613.

41. Azad,N., Zahnow,C.A., Rudin,C.M. and Baylin,S.B. (2013) The future of epigenetic therapy in solid tumours–lessons from the past. *Nat. Rev. Clin. Oncol.*, **10**, 256–266.

42. Altucci,L. and Rots,M.G. (2016) Epigenetic drugs: from chemistry via biology to medicine and back. *Clin Epigenetics*, **8**, 56.

43. Kelly,T.K., De Carvalho,D.D. and Jones,P.A. (2010) Epigenetic modifications as therapeutic targets. *Nat. Biotechnol.*, **28**, 1069–1078.

44. Chen,Z., Li,S., Subramaniam,S., Shyy,J.Y.-.J. and Chien,S. (2017) Epigenetic regulation: a new frontier for biomedical engineers. *Annu. Rev. Biomed. Eng.*, **19**, 195–219.

45. Falahi,F., Sgro,A. and Blancafort,P. (2015) Epigenome engineering in cancer: fairytale or a realistic path to the clinic? *Front. Oncol.*, **5**, 22.

46. Mlambo,T., Nitsch,S., Hildenbeutel,M., Romito,M., Müller,M., Bossen,C., Diederichs,S., Cornu,T.I., Cathomen,T. and Mussolino,C. (2018) Designer epigenome modifiers enable robust and sustained gene silencing in clinically relevant human cells. *Nucleic Acids Res.*, **46**, 4456–4468.

47. Shaffer,S.M., Dunagin,M.C., Torborg,S.R., Torre,E.A., Emert,B., Krepler,C., Beqiri,M., Sproesser,K., Brafford,P.A., Xiao,M. *et al.* (2017) Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*, **546**, 431–435.

48. Bailey,T., Krajewski,P., Ladunga,I., Lefebvre,C., Li,Q., Liu,T., Madrigal,P., Taslim,C. and Zhang,J. (2013) Practical guidelines for the comprehensive analysis of ChIP-seq data. *PLoS Comput. Biol.*, **9**, e1003326.

49. Karlić,R., Chung,H.-.R., Lasserre,J., Vlahovicek,K. and Vingron,M. (2010) Histone modification levels are predictive for gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 2926–2931.

50. van Galen,P., Viny,A.D., Ram,O., Ryan,R.J.H., Cotton,M.J., Donohue,L., Sievers,C., Drier,Y., Liau,B.B., Gillespie,S.M. *et al.* (2016) A multiplexed system for quantitative comparisons of chromatin landscapes. *Mol. Cell*, **61**, 170–180.

51. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

52. Gel,B. and Serra,E. (2017) karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics*, **33**, 3088–3090.

53. Kuhn,R.M., Haussler,D. and Kent,W.J. (2013) The UCSC genome browser and associated tools. *Brief. Bioinform.*, **14**, 144–161.

54. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.

55. Lam,A.J., St-Pierre,F., Gong,Y., Marshall,J.D., Cranfill,P.J., Baird,M.A., McKeown,M.R., Wiedenmann,J., Davidson,M.W., Schnitzer,M.J. *et al.* (2012) Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods*, **9**, 1005–1012.

56. Qu,H. and Fang,X. (2013) A brief review on the Human Encyclopedia of DNA Elements (ENCODE) project. *Genomics Proteomics Bioinform.*, **11**, 135–141.

57. Shcherbakova,D.M. and Verkhusha,V.V. (2013) Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat. Methods*, **10**, 751–754.

58. Druley,T.E., Vallania,F.L.M., Wegner,D.J., Varley,K.E., Knowles,O.L., Bonds,J.A., Robison,S.W., Doniger,S.W., Hamvas,A., Cole,F.S. *et al.* (2009) Quantification of rare allelic variants from pooled genomic DNA. *Nat. Methods*, **6**, 263–265.

59. Calvo,S.E., Tucker,E.J., Compton,A.G., Kirby,D.M., Crawford,G., Burtt,N.P., Rivas,M., Guiducci,C., Bruno,D.L., Goldberger,O.A. *et al.* (2010) High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. *Nat. Genet.*, **42**, 851–858.

60. Ezquerra-Inchausti,M., Anasagasti,A., Barandika,O., Garai-Aramburu,G., Galdós,M., López de Munain,A., Irigoyen,C. and Ruiz-Ederra,J. (2018) A new approach based on targeted pooled DNA sequencing identifies novel mutations in patients with Inherited Retinal Dystrophies. *Sci. Rep.*, **8**, 15457.

61. Patterson,N. and Gabriel,S. (2009) Combinatorics and next-generation sequencing. *Nat. Biotechnol.*, **27**, 826–827.

62. Erlich,Y., Chang,K., Gordon,A., Ronen,R., Navon,O., Rooks,M. and Hannon,G.J. (2009) DNA Sudoku–harnessing high-throughput sequencing for multiplexed specimen analysis. *Genome Res.*, **19**, 1243–1253.

63. Cao,C.-C. and Sun,X. (2016) Combinatorial pooled sequencing: experiment design and decoding. *Quant. Biol.*, **4**, 36–46.

64. Xu,W.S., Parmigiani,R.B. and Marks,P.A. (2007) Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene*, **26**, 5541–5552.

65. Yoshida,M., Horinouchi,S. and Beppu,T. (1995) Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays*, **17**, 423–430.

66. Komatsu,N., Kawamata,N., Takeuchi,S., Yin,D., Chien,W., Miller,C.W. and Koeffler,H.P. (2006) SAHA, a HDAC inhibitor, has profound anti-growth activity against non-small cell lung cancer cells. *Oncol. Rep.*, **15**, 187–191.

67. Wang,L., Zou,X., Berger,A.D., Twiss,C., Peng,Y., Li,Y., Chiu,J., Guo,H., Satagopan,J., Wilton,A. *et al.* (2009) Increased expression of histone deacetylaces (HDACs) and inhibition of prostate cancer growth and invasion by HDAC inhibitor SAHA. *Am. J. Transl. Res.*, **1**, 62–71.

68. Marks,P.A. and Jiang,X. (2005) Histone deacetylase inhibitors in programmed cell death and cancer therapy. *Cell Cycle*, **4**, 549–551.

69. Filippakopoulos,P., Qi,J., Picaud,S., Shen,Y., Smith,W.B., Fedorov,O., Morse,E.M., Keates,T., Hickman,T.T., Felletar,I. *et al.* (2010) Selective inhibition of BET bromodomains. *Nature*, **468**, 1067–1073.

70. Shi,J. and Vakoc,C.R. (2014) The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol. Cell*, **54**, 728–736.

71. Delmore,J.E., Issa,G.C., Lemieux,M.E., Rahl,P.B., Shi,J., Jacobs,H.M., Kastritis,E., Gilpatrick,T., Paranal,R.M., Qi,J. *et al.* (2011) BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*, **146**, 904–917.

72. Bid,H.K., Phelps,D.A., Xaio,L., Guttridge,D.C., Lin,J., London,C., Baker,L.H., Mo,X. and Houghton,P.J. (2016) The bromodomain BET inhibitor JQ1 suppresses tumor angiogenesis in models of childhood sarcoma. *Mol. Cancer Ther.*, **15**, 1018–1028.

73. Costa,D.D., Da Costa,D., Agathanggelou,A., Perry,T., Weston,V., Petermann,E., Oldreive,C., Wei,W., Stewart,G., Longman,J. *et al.* (2013) BET inhibition as a single or combined therapeutic approach in primary paediatric B-precursor acute lymphoblastic leukaemia. *Blood Cancer J.*, **3**, e126.

74. Christman,J.K. (2002) 5-Azacytidine and 5-aza-2′-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene*, **21**, 5483–5495.

75. Kaminskas,E. (2005) FDA drug approval summary: Azacitidine (5-azacytidine, VidazaTM) for injectable suspension. *Oncologist*, **10**, 176–182.

76. Corrales,M., Rosado,A., Cortini,R., van Arensbergen,J., van Steensel,B. and Filion,G.J. (2017) Clustering of Drosophila housekeeping promoters facilitates their expression. *Genome Res.*, **27**, 1153–1161.

77. Draker,R., Ng,M.K., Sarcinella,E., Ignatchenko,V., Kislinger,T. and Cheung,P. (2012) A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genet.*, **8**, e1003047.

78. Vardabasso,C., Gaspar-Maia,A., Hasson,D., Pünzeler,S., Valle-Garcia,D., Straub,T., Keilhauer,E.C., Strub,T., Dong,J., Panda,T. *et al.* (2015) Histone Variant H2a.Z.2 mediates proliferation and drug sensitivity of malignant melanoma. *Mol. Cell*, **59**, 75–88.

79. Liang,G., Gonzales,F.A., Jones,P.A., Orntoft,T.F. and Thykjaer,T. (2002) Analysis of gene induction in human fibroblasts and bladder cancer cells exposed to the methylation inhibitor 5-aza-2′-deoxycytidine. *Cancer Res.*, **62**, 961–966.

80. Schmelz,K., Sattler,N., Wagner,M., Lübbert,M., Dörken,B. and Tamm,I. (2005) Induction of gene expression by 5-Aza-2′-deoxycytidine in acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) but not epithelial cells by DNA-methylation-dependent and -independent mechanisms. *Leukemia*, **19**, 103–111.

81. Shen,H. and Laird,P.W. (2012) In epigenetic therapy, less is more. *Cell Stem Cell*, **10**, 353–354.

82. Miousse,I.R., Murphy,L.A., Lin,H., Schisler,M.R., Sun,J., Chalbot,M.-.C.G., Sura,R., Johnson,K., LeBaron,M.J., Kavouras,I.G. *et al.* (2017) Dose-response analysis of epigenetic, metabolic, and apical endpoints after short-term exposure to experimental hepatotoxicants. *Food Chem. Toxicol.*, **109**, 690–702.

83. Fardi,M., Solali,S. and Farshdousti Hagh,M. (2018) Epigenetic mechanisms as a new approach in cancer treatment: an updated review. *Genes Dis.*, **5**, 304–311.

84. Iorio,F., Knijnenburg,T.A., Vis,D.J., Bignell,G.R., Menden,M.P., Schubert,M., Aben,N., Gonçalves,E., Barthorpe,S., Lightfoot,H. *et al.* (2016) A landscape of pharmacogenomic interactions in cancer. *Cell*, **166**, 740–754.

85. Stratikopoulos,E.E., Dendy,M., Szabolcs,M., Khaykin,A.J., Lefebvre,C., Zhou,M.-.M. and Parsons,R. (2015) Kinase and BET inhibitors together clamp inhibition of PI3K signaling and overcome resistance to therapy. *Cancer Cell*, **27**, 837–851.