Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Biomedical Engineering

Advancing towards a global mammalian gene regulation model through single-cell analysis and synthetic biology

Josh Tycko[a,1], Mike V. Van[b,1], Michael B. Elowitz[c,d] and Lacramioara Bintu[e]

Abstract

Engineering complex genetic functions in mammalian cells will require predictive models of gene regulation. Since gene expression is stochastic, leading to cell-to-cell heterogeneity, these models depend on single-cell measurements. Here, we summarize recent microscopy and sequencing-based single-cell measurements of transcription and its chromatin-based regulation. Then, we describe synthetic biology methods for manipulating chromatin, and highlight how they could be coupled to single-cell measurements. We discuss theoretical models that connect some chromatin inputs to transcriptional outputs. Finally, we point out the connections between the models that would allow us to integrate them into one global input–output gene regulatory function.

Addresses

[a] Department of Genetics, Stanford University, Stanford, CA 94305, USA

[b] Department of Biology, Stanford University, Stanford, CA 94305, USA

[c] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[d] Howard Hughes Medical Institute (HHMI), Department of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA

[e] Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

Corresponding author: Bintu, Lacramioara (lbintu@stanford.edu)

[1] These authors contributed equally to this work and are listed in alphabetical order.

Current Opinion in Biomedical Engineering 2017, **4**:174–193

This review comes from a themed issue on **Synthetic Biology and Biomedical Engineering**

Edited by **Charlie Gersbach**

Received 25 July 2017, revised 4 October 2017, accepted 26 October 2017

https://doi.org/10.1016/j.cobme.2017.10.011

Keywords

Gene regulation, Chromatin, Single-cell, Modelling, Synthetic biology.

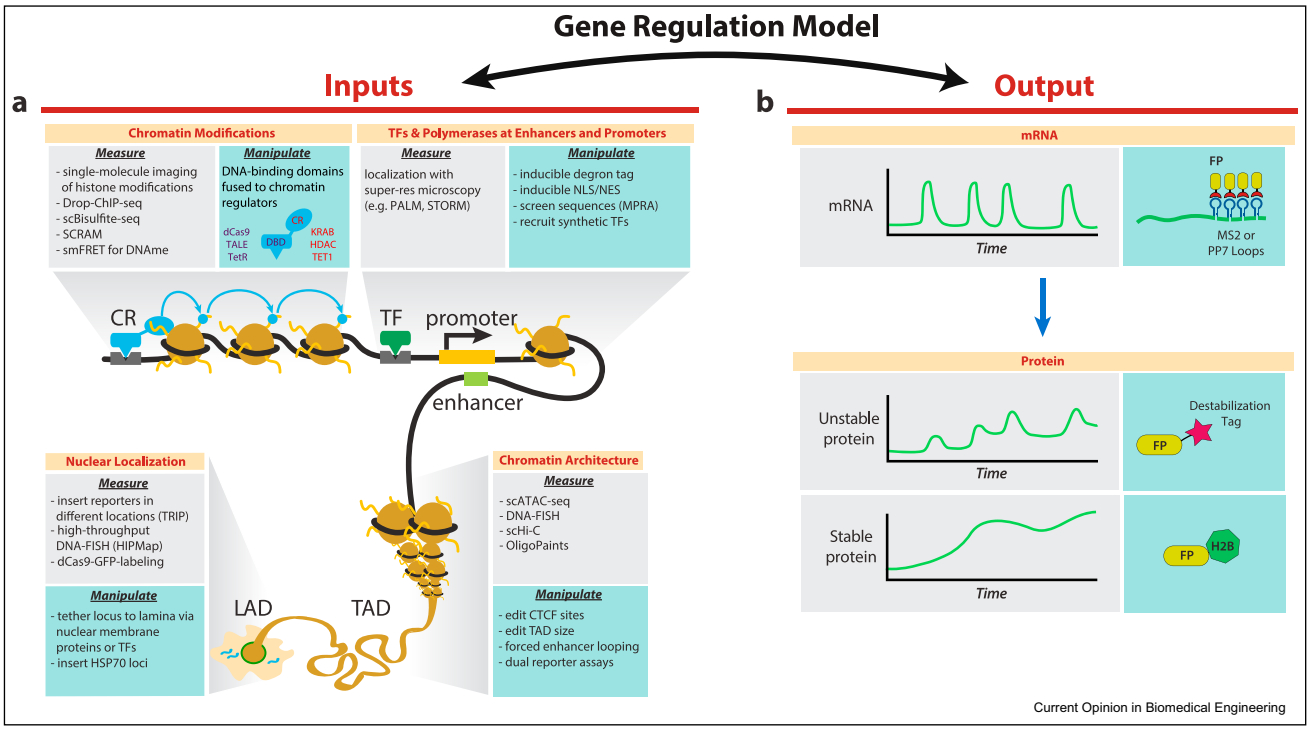
Introduction

Transformative increases in our understanding of gene regulation, along with technologies like sequencing and genome editing, allow us to envision a new era of medicine, in which diagnosis is personalized and therapies are targeted to specific genes, through manipulations of their expression patterns and epigenetic states. One key obstacle to realizing this vision is that we currently lack a predictive, quantitative, and unified theoretical framework for eukaryotic gene regulation. Such a framework would provide an effective "input–output" response for each gene, representing the dependence of its expression level and epigenetic memory on inputs like transcription factor concentrations, specific epigenetic marks, and location within the nucleus. Constructing such a framework is challenging because it has to incorporate many distinct mechanisms of gene regulation.

To experimentally map the effective input–output behavior of a gene, two types of technologies are essential. First, because gene regulation is a highly stochastic process, it can only be understood quantitatively through single-cell measurements. Second, the ability to design synthetic genes and alter natural ones is critical for exploring how sequence, chromatin state, cell context, and specific regulatory factors collectively control gene expression. Mammalian single-cell analysis and synthetic biology are rapidly becoming much more powerful, making this an ideal time to address these problems.

Here we review recent work that has taken promising initial steps towards the kind of quantitative, single-cell gene regulation framework that will enable us to understand, predict, and design gene regulatory systems in mammalian cells. We highlight recent developments in measuring the chromatin inputs (Fig. 1a) and transcriptional output (Fig. 1b) using single-cell methods (Table 1), manipulating those inputs with synthetic biology (Table 2), and modeling the input–output function (Fig. 2). We omit many important factors and mechanisms controlling gene expression including: splicing, non-coding RNAs, post-transcriptional regulation (e.g. miRNA), RNA stability and export, translational and post-translational protein control. However, all of these will ultimately need to be incorporated for a more comprehensive description of gene regulation.

Fig. 1



Input-output function of a unified single-cell model of gene regulation. To develop a complete model of gene regulation, a function that relates inputs like transcription factor binding and chromatin state to transcriptional output is required. **(a)** *Inputs: modulators of transcription, single-cell measurements and synthetic manipulation.* Recently developed tools to measure (gray panels) and manipulate (cyan panels) many determinants of gene regulation (i.e. chromatin modifications, transcription factors and polymerases at enhancers and promoters, chromatin architecture, and nuclear localization). *Abbreviations:* ***scBisulfite-seq****: single cell bisulfite sequencing of DNA methylation;* ***smFRET for DNAme****: single-molecule fluorescence resonance energy transfer measurement of DNA methylation variants;* ***SCRAM***: single-cell restriction analysis of methylation; ***Drop-ChIP-seq****: chromatin immunoprecipitation in droplets followed by sequencing;* ***DBD****: DNA-binding domain;* ***CR****: chromatin regulator;* ***TF****: transcription factor;* ***NLS/NES****: nuclear localization/export signal;* ***MPRA****: massively parallel reporter assay;* ***scATAC-seq****: single cell assay for transposase accessible chromatin followed by sequencing;* ***scHi-C****: single cell high-throughput chromosome conformation capture;* ***DNA-FISH****: DNA fluorescent in situ hybridization;* ***CTCF****: CCCTC-binding factor;* ***TAD:*** *topologically associated domain;* ***LAD****: lamina associated domain;* ***TRIP****: thousands of reporters integrated in parallel.* **(b)** *Output: dynamic measurements of transcription at the single-cell level.* Transcription from a specific locus can be monitored by analyzing fluorescently labeled mRNA over time (top panel) [137,166,167]. The RNA is genetically modified to contain MS2 or PP7 RNA loops that are recognized by cognate RNA-binding proteins fused to fluorescent proteins (FP). Transcription activity can also be inferred from the levels of fluorescently labeled proteins over time (bottom panel). For proteins fused to a destabilization tag, the levels of fluorescence mirror the changes in mRNA. For stable proteins (fused to H2B), the slope of fluorescence intensity (change per unit time) reflects the level of transcription [23,168].

Output: measurements of transcription at the single-cell level

Here, we define the output of gene regulation as the level of mRNA produced, as well as its variability over time and from cell to cell. Work in both cell culture and embryos has revealed that transcription occurs in bursts, alternating between 'on' and 'off' periods of mRNA production (reviewed in Refs. [1–4]). Mammalian mRNA bursting was indirectly inferred by analyzing mRNA distributions in fixed single cells by microscopy using Fluorescence *in Situ* Hybridization (RNA-FISH) [5]. Theoretically, burst-like gene expression processes produce a negative binomial distribution of mRNA numbers, different from the Poisson distribution that would be expected if mRNA were produced at a constant rate. This difference is especially important when the rate of bursting is low, leading to high cell-to-cell variability in the form of a long tail of cells with large numbers of mRNAs [2]). RNA-FISH has been recently modified to measure hundreds to thousands of mRNA species in single cells by using sequential rounds of hybridization and removal of the fluorescent signal, allowing for "barcoding" each mRNA species [6,7]. These types of multiplexed measurements, including single-cell RNA sequencing methods [8,9], can be used to measure the output of all genes that share a common input in single cells. They can be especially powerful when used as an end-point measurement coupled to a time-course recording of the dynamic history of gene expression (movie-FISH) [10,11].

Table 1

Single-cell measurements of gene regulatory inputs and outputs.

Category	Method	Measures	Advantages	Limitations	Ref.
Microscopy	RNA-FISH: Fluorescent *in situ* hybridization of selected RNA	RNA	- Quantification and localization of individual mRNA molecules	- Samples need to be fixed - Not dynamic - Limited to 1–4 mRNAs per cell (number of available fluorophores)	[5]
	seqFISH: sequential barcoded FISH	**RNA**	- Quantification and localization of tens-hundreds of RNA species in single cells - Sequential hybridization and removal of fluorescent probes - scales exponentially with the number of fluorophores	- Not dynamic - Requires specialized synthesis of probe library - Problems quantifying high abundance and aggregated transcripts (by imaging)	**[6]**
	MERFISH: multiplexed error-robust FISH	RNA	- Quantification and localization of hundreds-thousands of RNA species in single cells - Combinatorial hybridization labeling and encoding scheme for error correction	- Not dynamic - Requires specialized synthesis of probe library - Problems quantifying high abundance and aggregated transcripts (by imaging)	[7]
	FISSEQ: Fluorescent *in situ* sequencing	**RNA**	- Quantification and localization of hundreds of RNA species in single cells	- Not dynamic - Can't detect low abundance transcripts, limited by mRNA to cDNA conversion - Low signal-noise ratio between bound and unbound fluorophore	**[169]**
	RNA binding protein-fluorophore fusions that bind RNA stem loops (e.g. MS2,PP7)	RNA	- Quantification of nascent mRNA - Records temporal dynamics	- Need to engineer the mRNA - Stem loops may affect the stability of the mRNA	[12,13]
	RNA Mimics of GFP (e.g. spinach, broccoli)	**RNA**	- Quantification of mRNA levels - Records temporal dynamics	- Need to engineer the mRNA - Aptamer may affect the stability of the mRNA - Signal of original aptamers was low	**[14]**
	RNA-targeting Cas-fluorophore	RNA	- Measures localization of concentrated RNA (e.g. stress granules) - Can target any RNA without genetically engineering it	- Low signal-noise ratio due to unbound Cas-fluorophore	[15,16]
	Super-resolution microscopy (e.g. PALM, STORM) of fluorescently labeled TFs, polymerases, or histone modifications	**Proteins**	- Track the locations of single proteins - Temporal dynamics	- Photobleaching of fluorophore (duration imagining limited) - Phototoxicity	**[20,22,27–30]**
	Inserting methylation-sensitive fluorescent reporter near CpG region of interest	5mC (indirectly)	- Live single-cell fluorescent visualization of changes in DNA methylation states at a specific locus	- Other silencing factors may affect the reporter	[81]
	Single molecule fluorescence energy transfer (smFRET) of DNA methylation and hydroxymethylation	**5mC and 5hmC**	- Able to detect and quantify 5hmC and 5mC single molecules from trace amounts of DNA - Detects multiple DNA modifications on the same DNA	- 5mC labeling relies on Tet1 conversion to 5hmC; Tet1 is inefficient - Incomplete labeling - Photobleaching of fluorophores - Not dynamic	**[86]**

	DNA-FISH: Fluorescent *in situ* hybridization of selected DNA in fixed cells	3D-chromatin and nuclear localization	- Localization of individual DNA molecules within the nucleus	- Not dynamic - Limited number of loci	[170] comparison with Hi-C
	OligoPaints	**3D-chromatin**	- When combined with super resolution microscopy, it can reveal 3D chromatin organization in a single cell	- Single locus (or limited number) - Requires specialized synthesis of probe library	**[116,117]**
	HIPMap: High-throughput imaging method for mapping spatial gene positions	Nuclear localization	- Automated, high throughput imaging and analysis pipeline for 3D analysis of DNA-FISH probes	- Automation of image analysis may detect false FISH signals	[141]
	dCas9-fluorophore labeling of genomic loci	**Nuclear localization**	- Live visualization of endogenous loci - 3D visualization	- Potential off-target binding of dCas9-fluorophore - Low signal-noise ratio between bound and unbound fluorophore (better at repetitive sequences or with multiple targeting)	**[142]**
Sequencing	Drop-ChIP-seq: Chromatin immunoprecipitation from single cells isolated in droplets	Histone modifications	- High-throughput analysis of chromatin states in single cells using microfluidics and DNA barcoding	- Sparse coverage of the genome in the single cells	[66]
	scWGBS: Whole-genome sequencing of DNA methylation by bisulfite conversion of sorted single cells	**5mC or 5hmC, indistinguishably**	- Whole-genome - Single-base resolution of CpG and non-CpG methylation sites	- Bisulfite conversion cannot distinguish between 5mC and 5hmC - Allele-specific differences in methylation are difficult to detect - Sparse coverage - DNA degradation occurs during bisulfite conversions	**[82]**
	SCRAM: Single-cell restriction analysis of methylation	5mC	- No DNA degradation; thus, no locus dropout - Accurate analyses of CpG methylation at a specific genomic locus in single cells	- Limited number of loci - Readout is a binary (yes or no) for the entire methylation state of a locus	[83]
	sc-GEM: Single-cell analysis of genotype, expression and methylation	**DNA RNA 5mC**	- Simultaneous measurement of genomic DNA variant, gene expression, and DNA methylation - Single-cell (using microfluidics) - no locus dropout	- Only a snapshot, not dynamic - Limited number of loci, not whole-genome - Readout is a binary (yes or no) for the entire methylation state of a locus	**[171]**
	scAba-seq: Single-cell 5hmc sequencing by 5hmC glucosylation and glucosylation-dependent digestion of DNA by AbaSI	5hmC	- Highly specific, genome-wide detection of the 5hmC epigenetic mark	- Limited coverage in single cells	[85]
	scATAC-seq: Single cell assay for transposase accessible chromatin with sequencing	**3D-chromatin accessibility**	- Simple library preparation protocol - Requires no sonication, antibodies, or sensitive enzymatic digestions - Can identify co-accessible regions (enhancers and promoters)	- Snapshot, not dynamic - Limited coverage in single cells	**[101–104]**

(*continued on next page*)

Table 1 (*continued*)

Category	Method	Measures	Advantages	Limitations	Ref.
	scHi-C: High-throughput chromosome conformation capture in single cells	3D-chromatin contacts	- Segregate 3D genome maps by where they are in the cell cycle	- Sparse 3D map and limited coverage - Require a large amount of sequencing	[107–110]
	TRIP: Thousands of Reporters integrated in Parallel	**Gene expression**	- High throughput and multiplexed - Do not need to isolate clonal cell lines because reporters are barcoded	- Random integration of reporters into essential genes could cause cell death	**[140]**
	MPRA: Massively Parallel Reporter Assay	Gene expression	- High-throughput - Lentiviral constructs can accommodate large inserts (up to 10 kb), can assay long regulatory sequences - Lentivirally-delivered reporter genes display relatively fast expression dynamics (<48 h)	- Oligo synthesis limits size of regulatory sequence that can be tested	[53,54]

Such temporal dynamics of a particular gene can be directly visualized in live cells by detecting fluorescently-labeled mRNA molecules at the transcription initiation site (Fig. 1b) [12–14]. New CRISPR-based methods could make it possible to track any mRNA of interest by reprogramming RCas9 [15] or Cas13a [16] with guide RNAs, doing away with the need to engineer the mRNA itself to include label-binding sites.

Transcription dynamics can also be inferred from measuring protein expression. However, the impact of transcription on protein levels is sensitive to mRNA and protein degradation rates (Fig. 1b). When the reporter protein and its transcript are destabilized, fluctuations in protein levels track transcriptional burst kinetics [17]. At the other extreme, for very stable proteins such as histone 2B (H2B)-fluorescent protein fusions, the derivative of protein accumulation over time is a good proxy of promoter activity [18,19]. These live-cell experiments, measuring protein or mRNA outputs, allow readout of the transcriptional output over time, in individual cells, during and after modulation of a transcriptional input.

Transcriptional variation occurs across many timescales, which can relate to the genomic length scale of the causative molecular changes. For example, fast bursting dynamics (seconds-minutes) can result from transcription factors (TFs) binding, RNA polymerase (RNAP) binding and initiating transcription, or chromatin modifications with fast turnover rates (i.e. histone acetylation) [20–22]. Longer time-scale switching (hours-days) can result from changes in more stable chromatin modifications such as DNA and histone methylation, which are associated with epigenetic memory [11,23,24]. Therefore, in order to understand transcription at different time scales, we need to systematically measure and manipulate its inputs across different chromatin length scales.

Transcriptional inputs: single-cell measurements and synthetic manipulation

Transcription from a specific DNA locus is controlled by multiple factors (Fig. 1a), including RNA polymerase and transcription factors (TFs) bound at promoter and enhancer sequences, nucleosome occupancy, chromatin modifications and regulators, the 3D structure of chromatin, and nuclear localization (reviewed in Ref. [1]). In this section, we review the latest advances that allow us to measure each of these transcriptional control layers at the single-cell level and manipulate them using synthetic biology approaches (reviewed in Refs. [25,26]).

Polymerase and transcription factors at promoters and enhancers

In order to quantify the effect of TFs and polymerases on transcription, we need to measure and manipulate

their concentration, DNA binding kinetics, activation/repression strength, and binding site number and position.

Measurements

Advances in super-resolution microscopy and fluorescent labeling allow monitoring of single polymerase and TF molecules over time in the nucleus (reviewed in Ref. [27]). Simultaneous imaging of both RNA polymerase II (Pol II) binding at a promoter (by PALM) and the mRNA produced from that locus (by STORM) revealed that RNA polymerases assemble in transient clusters (of ~80 polymerases) that last 5–25 s. The lifetime of Pol II clusters correlates linearly with the number of mRNAs produced during that burst of transcription [20,21]. A subset of the TFs tracked with super-resolution microscopy were also shown to form dynamic clusters (e.g. Sox2 [28]) or to diffuse in a compacted manner suggestive of local interactions with nuclear structures (P-TEFb [29]). Indeed, the Sox2 clusters co-localized with Pol II-rich regions [30], suggesting a direct relationship between local TF and polymerase concentrations and transcriptional bursting. The exact molecular mechanisms of Pol II and TF dynamic clustering and their connection with the 3D chromatin architecture are not yet known. However, it is likely that these dynamic clusters rely on strong binding of TFs to chromatin (Sox2) and synergistic interactions between TFs and the Pol II CTD tail (Sox2 and P-TEFb) [30], as TFs that don't have these features diffuse more freely around the nucleus (c-Myc [29]). Regardless of the mechanism, this information about the dynamic local distributions of TFs in the nucleus can now be incorporated into theoretical models of gene expression [31].

Manipulations

The ideal way to determine gene expression input–output relationships would be to precisely control the concentration and binding dynamics of TFs at promoters and enhancers and measure their effects on expression. Modulating the concentration of endogenous TFs at a gene is generally done by allowing cells to differentiate into cell fates with distinct TF activity profiles, or by stimulating them with a global signal (e.g. serum [32,33], Zinc [32], or TNF-alpha [34]). While these stimuli resemble physiological perturbations, they can lead to off-target changes in other factors that may influence transcription, such as inducing other TFs and changing local or global chromatin state.

In contrast to these broad perturbations, modern synthetic tools enable faster and more precise control over TF concentration in the nucleus. Simply turning on/off expression of a gene encoding for the TF is not ideal because it's slow to respond (hours) and can be noisy, especially at low expression levels (due to transcriptional bursting). Faster and more graded regulation of TF availability can be achieved by controlling mRNA/protein stability or nuclear localization by fusing the mRNA or protein to a tag responding to external signals that don't affect the rest of the cell (reviewed in Refs. [26,35]). These signals can be: 1) chemical small molecules that can easily diffuse inside cells and control mRNA stability via ribozymes (reviewed in Ref. [36]), protein stability via degradation domains [37–39], or protein localization [40–42]; 2) optical specific wavelengths of light that control protein nuclear localization [43–49] or stability [50,51]; and more recently 3) magnetic or radio waves [52]. With the advent of CRISPR technology and viral delivery methods, these types of tags are now easier to install on various TFs in order to systematically and precisely sweep through a wide range of input TF concentrations.

Beyond concentration, one can vary the TF binding site strengths and numbers by changing the DNA sequence of a promoter or enhancer. DNA synthesis and delivery methods have made it easier to generate barcoded libraries that test tens of thousands of DNA constructs bearing TF binding motifs, in massively parallel reporter assays [53,54]. These high-throughput methods effectively sample a large sequence space, provide direct measurements of DNA sequence as an input to transcription, and infer combinations of TFs that act synergistically.

Alternatively, synthetic TFs with tunable DNA binding and activation/repression domains can be used to manipulate transcriptional outputs. While TFs with inducible bacterially-derived DNA binding domains, such as TetR-VP16, proved useful for early studies of transcriptional bursting [5], more recent programmable DNA binding domains (ZFs, TALEs and dCas9) allow for multiplexing and ease of exploration of binding site location and number at endogenous genes [32,55–58]. The number and strength of activation or repression domains can also be tuned for synthetic TFs, making them an appealing tool for dissecting the input–output transcriptional regulatory function. Overall, these approaches to tune the concentration of TFs, alter the DNA sequence of promoters or enhancers, and target synthetic TFs are enabling precise manipulation of gene regulation directly at the site of transcription initiation.

Chromatin modifications and chromatin regulators

Polymerase and TF binding is strongly correlated with chromatin modifications [59–61], as has been acknowledged since the discovery of chromatin modifications [62]. However, improving our understanding from correlations to causation remains a major challenge. Recent technological advances in imaging, sequencing,

Table 2

Synthetic biology manipulations of gene regulatory inputs.

Method	Manipulated Input	Category	Purpose	Limitations and Improvements	References
RNA-based gene control with aptamers	TF concentration	RNA engineering	Tune concentration of TFs with a small molecule-responsive ribozyme integrated in the TFs' mRNA	- Cannot usually go to zero RNA levels in mammalian cells - Small molecule may interfere with other cell functions	[36]
Inducible NLS/NES		Protein engineering	An amino acid sequence that tags a protein for import into (NLS) or export out (NES) of the nucleus	- Intrinsic NLS of TF may interfere with localization - Having multiple NLS/NES can improve efficiency	[40–42]
Inducible degron tag		Protein engineering	Fusing TF with degron tag enables controlled degradation	- Fusing TF with degron may disrupt normal TF function	[37–39]
Recruit TFs with DNA-binding domains	**TF localization**	**Protein engineering**	**Recruitment of TFs to promoters or enhancers via programmable DBDs (i.e. ZFs, TALEs, dCas9)**	- Binding strength and efficiency of DBDs can vary, confounding results	**[32,58,172,173]**
Recruit CRs with DNA-binding domains	**CR localization**	**Protein engineering**	**Fusion to DBDs enables recruitment of CRs to specific genomic sites to modify chromatin**	- Binding strength and efficiency of DBDs can vary - CRs are generally large, which can make delivery inefficient	**[48,87,92–94,98,172,174–181]**
Insert HSP70 loci near gene of interest	Gene localization	Genome engineering	Inserting HSP70 loci near endogenous genes enables heat-shock inducible control of gene localization to the nuclear interior	- Inserting the HSP70 loci into the genome with CRISPR-Cas9 will allow for greater flexibility in sites of insertion	[148]
Tether locus to lamina via nuclear membrane protein		Protein engineering	Recruiting nuclear membrane protein to locus (via LacI) allows for tethering of gene to nuclear periphery	- Recruiting nuclear membrane proteins with dCas9 will allow for programmable nuclear periphery localization of any gene	[144–147]
Remove, insert, or flip CTCF binding sites	**Chromatin organization**	**Genome engineering**	**Small changes in genome sequence enable chromatin re-organization of large domains**	- CRISPR-Cas9 insertions and paired-guide deletions are relatively inefficient - dCas9-targeted DNA methylation may be a more efficient means of inhibiting CTCF binding	**[98,121–123]**
Edit TAD size with *Cre-loxP* recombineering		**Genome engineering**	**Grow and shrink TADs by 100's of kilobases, or generate inversions and duplications of TADs**	- More efficient for large deletions and inversions but lower-throughput than recent CRISPR-Cas9 approaches	**[130]**
Forced enhancer looping		**Protein engineering**	**Fusing a domain of the nuclear factor Ldb1 to a designer zinc finger creates a synthetic protein that forces an enhancer-promoter loop at the β-globin locus**	- Different looping mechanism than most endogenous chromatin loops - Zinc fingers are relatively laborious to re-target to other loci - Alternative CRISPR-Cas9 CLouD platform is simpler to re-target, but has not yet been tested at many loci	**[135,136]**

and synthetic biology make it possible to measure and manipulate the level of chromatin regulators and modifications at a specific locus in single cells (reviewed in Refs. [63,64]).

Histone modifications measurements

In order to relate changes in chromatin modification to changes in transcription at a specific gene, one needs to quantitatively measure chromatin modifications in single cells at that locus. This is more difficult than measuring transcription in single cells, because both the chromatin modifications and the DNA sequence need to be detected at the same time, and the signal is weak (especially from chromatin, due to weak antibody binding to histone modifications).

One approach to solving this problem is to amplify the signal by inserting hundreds of copies of a gene at a genomic locus. Using this approach in conjunction with imaging via antigen-binding fragments (Fabs) conjugated to fluorescent dyes it was possible to monitor TF-GFP local concentration, histone acetylation, and RNA polymerase in live cells [22]. These data showed that histone acetylation speeds up the binding kinetics of TFs, as well as the RNAP transition from initiation to elongation. Another imaging approach is to amplify the co-localization signal between a gene labeled with a DNA biotin probe and an antibody against a histone modification by using a proximity ligation assay (PLA) in fixed cells [65]. This method could be used as an end-point measurement in conjunction with time-lapse microscopy analysis of gene expression to study the relationship between transcription and histone modifications.

An alternative approach to imaging is to separate the cells with microfluidics, barcode them, and then pool them and perform traditional chromatin immunoprecipitation followed by DNA sequencing (Drop-ChIP) [66]. These measurements revealed that chromatin modifications are indeed heterogeneous in embryonic stem cells, especially at pluripotency enhancers and Polycomb targets. In the future, this technique could be combined with single-cell RNA-seq to interrogate the relationship between chromatin and RNA levels, although currently data from both techniques are too sparse. More recently, single-molecule imaging of multiple histone modifications on nucleosomes immobilized on slides was combined with single-molecule sequencing of the DNA associated with each nucleosome [67]. This technique demonstrated that active and repressive modifications (H3K4me3 and H3K27me3) can co-exist on the same nucleosome and that the frequency of finding them together is higher in embryonic cells compared with differentiated cells. This finding underscores the need to understand how changing the cell type modulates the dynamics of chromatin-mediated gene regulation. Overall, these approaches will help decrypt the dynamic combinations of modifications that are read by transcription complexes in the single cells that make up heterogeneous populations and tissues.

DNA modification measurements

The main DNA modification known to modulate transcription is cytosine methylation (5mC) at CpG dinucleotides. However, over the last decade, new DNA modifications, such as 5-hydroxymethylcytosine (5hmC) [68,69] and N^6-methyladenine (N6-mA) [70], have surfaced as potential modulators of transcription in mammalian cells. The causal and quantitative effect of these modifications on gene expression is still under investigation. Since it is becoming clear that there is a lot of cell-to-cell variability in the level of these modifications, single-cell methods are essential in dissecting both the role of these modifications in development and disease, and their connection to transcription.

Historically, 5mC has been associated with repressed transcription (reviewed in Refs. [71,72]), and 5hmC was thought to simply serve as an intermediate in the process of methylation removal and, thus, was associated with gene activity (reviewed in Ref. [73]). However, this classical view is currently being revised. Since DNA methylation can directly modulate the binding of diverse TFs and organizational proteins (histones, CTCF) [74,75], its effects on transcription depend on the sites affected. Notably, there are important exceptions in which DNA methylation activates gene expression, for example, by eliminating CTCF-binding at an insulator domain separating a distal enhancer from its promoter [76,77]. Recently, methylation preferences for a large library of transcription factors were measured by methylation-sensitive SELEX (systematic evolution of ligands by exponential enrichment) [78]. This *in vitro* method revealed that many activators, including POU5F1 (OCT4) and HOXB13, preferentially bind motifs with 5mC, and this result was validated by ChIP in mouse embryonic stem cells (mESCs) and crystallography. Moreover, 5hmC can also modulate the binding of TFs and CRs [79,80] and, hence, could serve as an additional direct input to transcription, especially in stem cells and neurons where this modification is found in abundance. Our understanding of DNA methylation's impact on gene regulation is deepening as these mechanistic links are revealed, and its multiple functions underscore the need for methods that can simultaneously detect methylation and transcriptional dynamics.

In that direction, a methylation-sensitive fluorescent reporter was developed to monitor 5mC presence at a locus in single cells over time: it consists of a minimal imprinted gene promoter that can be repressed or activated depending on the DNA methylation state of adjacent sequences [81]. The reporter was inserted near super-enhancers associated with pluripotency (of *Sox2* and *mirR290*) to monitor loci-specific DNA methylation

changes during mouse ES cell differentiation and reprogramming of mouse fibroblast cells into iPSCs. This reporter could also be used to answer basic questions about how DNA methylation spreads and how the level of methylation quantitatively affects transcription. However, to answer these questions, one would also need to measure the levels of 5mC methylation in the same cells.

Multiple single-cell methods for detecting DNA methylation at a specific locus or genome-wide were developed recently (reviewed in Refs. [63,64]), although some of them can't differentiate between 5mC and 5hmC. Traditionally, DNA methylation was measured using bisulfite conversion, a method that converts unmodified cytosines (Cs) to uracils (Us). This process has been adapted to measure the lack of cytosine modifications in single cells in a method called whole genome bisulfite sequencing (scWGBS) [82]. However, since this technique cannot differentiate between 5mC and 5hmC, it is not ideal for dissecting their effects on gene expression. Alternatively, mapping of DNA methylation can be done with restriction enzymes that are inhibited by specific DNA modifications. For instance, single-cell restriction analysis of methylation (SCRAM) combines microfluidics, restriction digestion, and locus-specific DNA amplification [83]. While this technique was used with enzymes that cut unmodified Cs (and hence mapped both 5mC and 5hmC), in principle it could be modified to map 5hmC by using PvuRts1, an enzyme that specifically cuts 5hmC DNA [84]. A similar restriction-based approach was used in Aba-seq for single-cell genome-wide 5hmC mapping in which the enzyme, AbaSI, caused 5hmC glucosylation and glucosylation-dependent digestion of DNA [85]. This study revealed that two opposite strands of DNA can display up to tenfold difference in hydroxymethylation levels [85]. This strand bias may serve as a source of strand-specific (and hence cell-progeny specific) epigenetic memory to regulate DNA-binding protein interactions.

Intriguingly, 5mC and 5hmC can co-exists in very close proximity on DNA purified from mouse cerebellum, as measured by single-molecule fluorescence resonance energy transfer (smFRET) between fluorophore-labeled 5mC and 5hmC sites [86]. Moreover, hydroxymethylated sites on the opposite strand of DNA methylated sites inhibited the binding of repressive DNA-binding domains (e.g. methyl-CpG binding domains) at these sites; thus, suggesting a potential function of 5mC-5hmC combinations in promoting gene expression. Together, results from single-cell mapping of DNA modifications underscore the cell-to-cell heterogeneity of these modification, their importance in cell decisions during differentiation and reprogramming, as well as the need for assays that can distinguish between the different types of DNA modifications if we are to understand their effect on gene expression.

Manipulations of chromatin modifications

Synthetic biologists are building a toolbox for precision epigenome editing by recruiting different chromatin regulators along the genome (reviewed in Refs. [55,56,87,88]). Since in mammalian cells most chromatin regulators (CRs) do not have consensus DNA binding sites, they have been traditionally recruited to DNA by fusing them with DNA binding domains borrowed from bacteria or yeast (LexA [89], Gal4 [90], or TetR [91]) and introducing their cognate sequences in the genome, either near an artificial reporter or endogenous gene. Recently, this method has been used to shed new light on the effect of silencing chromatin regulators on chromatin spreading [24] and the dynamics of gene expression in single-cells [23,24]. These studies have shown that upon recruitment of various silencing CRs (HP1, EED, KRAB, DNMT3B, HDAC4), a fraction of cells silence completely in a stochastic manner, and this fraction depends on the recruitment time. Moreover, after the end of recruitment some cells retain permanent epigenetic memory, and the fraction of cells permanently silenced depends on the duration and strength of CR recruitment [23,24]. However, the ability to silence a gene permanently does not depend only on the CR recruited, but also on cell type [24,92–94] and local genomic context [93,94]. A quantitative and systematic analysis of how these variables interact to determine the dynamics of CR-mediated silencing, activation, and epigenetic memory would enable regulation of genetic circuits over wide domains of time and component concentrations.

Looking forward, programmable DNA-targeting tools (zinc fingers, TALEs and dCas9) enable a much more systematic and direct mapping of the connection between chromatin regulators and gene expression (reviewed in Refs. [55,56,88,95]. Fusing CRs to programmable DNA binding domains allows recruitment of various CRs to any genomic locus (without installing cognate sequences), including both promoters and distal enhancers [96,97]. For example, targeting of dCas9-Tet1 fusions to endogenous promoter or enhancer sequences has recently been utilized to demethylate these sequences; thus activating the expression of genes involved in neuron development or facilitating reprogramming of fibroblasts into myoblasts [98]. Combining these new tools with single-cell dynamic monitoring of transcription will greatly improve our understanding of chromatin-mediated transcriptional control of endogenous genes.

Dynamic chromatin accessibility

In the nucleus of a human cell, two meters of chromosomal DNA is compacted ~10,000 fold into chromatin. As such, gene expression depends on the controlled physical organization of the genome, ensuring that certain genes are accessible to transcription machinery while meeting strict spatial requirements.

Measurements

DNA accessibility has historically been measured by digesting chromatin with DNase I and determining which sequences were hypersensitive to cleavage, as this implies they were not protected by compacted histones [99,100]. More recently, accessibility has been measured in bulk and single cells with Assay for Transposase-Accessible Chromatin (ATAC-seq), in which an engineered Tn5 transposase selectively integrates sequencing adapters into accessible chromatin [101–104]. The single-cell ATAC-seq experiments demonstrate that cell-to-cell variability in accessibility is associated with specific *trans*-factors, particularly at binding motifs where multiple transcription factors compete (e.g. GATA1 and GATA2) or where chromatin effectors bind (e.g. BRG1 and P300) [102]. Improving scATAC-seq throughput with combinatorial indexing enabled 15,000 single-cells to be assayed in parallel [103], but the data's sparsity still presents challenges. Further development to couple these techniques with RNA measurements in the same cells, by using microscopy or sequencing, may clarify the relationship between accessibility variation and transcriptional variation. Intriguingly, local variation of accessibility is highly correlated with large-scale chromosome organization into compartments [102,105], which suggests that the larger 3D genome structure may help explain the variable accessibility of these short DNA sequences.

The 3D dynamic structure of chromatin

Beyond local accessibility of specific DNA sequences, transcription of many genes depends on their interaction with distal regulatory elements, such as enhancers [106]. Dynamic promoter–enhancer interactions, and the chromatin domains which structure those interactions, can be measured and manipulated to reveal their contributions to transcriptional outputs.

Measurements

Long-range chromatin interactions and loops are predominantly measured with variations on two foundational approaches: 1) DNA Fluorescent *in Situ* Hybridization (DNA-FISH) and 2) chromosome conformation capture-based methods (3C). DNA-FISH imaging provides the distribution of the spatial distance between two selected genomic locations across many single cells. High-throughput 3C-based methods (Hi-C) measure the frequency of contacts between all pairs of genomic locations by cross-linking chromatin, digesting the DNA into short fragments, and then performing a dilute 'proximity ligation' reaction followed by sequencing across the ligation junctions [107–110]. At the megabase scale, Hi-C shows the genome is organized into highly-interacting Topologically Associated Domains (TADs), which rarely interact across domain boundaries [108,111].

Single-cell Hi-C data demonstrates heterogeneity in contact frequencies [112,113], which could contribute to variation in gene expression. These single-cell contact maps can be phased in cell-cycle pseudotime to study and model how loops, TADs, and larger chromatin compartments stabilize and dissolve [114]. These single-cell data should be useful for constraining models of chromosome conformation to fit their measured variation, and thus represent an important step towards a model of transcription that can account for the dynamic 3D genome. However, as a static snapshot method, Hi-C lacks the ability to explain the dynamics of individual loops that affect expression of particular genes.

Ultimately, live-cell microscopy could measure these dynamics. Already, static FISH is established as an important complement to Hi-C, because it directly measures spatial distances in intact single cells. Recently, the integration of FISH with super resolution microscopy (STORM) has revealed distinct properties of chromatin domains in different epigenetic states, with unprecedented resolution [115]. Entire domains in the active, inactive, or polycomb-repressed states were visualized by hybridizing FISH probes across kilobase-to-megabase regions [116,117]. This led to the discovery that chromatin domains obey a power-law scaling between their length and physical volume, with different scaling exponents depending on the chromatin state. Notably, polycomb-repressed chromatin is the densest, and its density increases with domain length.

A diverse, growing suite of chromatin architecture measurements [110,118,119] should uncover key parameters for the transcriptional input–output function, including enhancer-promoter interaction frequencies, and structure of insulated domains that constrain chromatin state spreading. Coupling these chromatin structure measurements with single-cell expression measurements could reveal the relationship between chromatin architecture inputs and transcriptional outputs.

Manipulations

Manipulating the size and organization of chromatin domains, or TADs, also presents a path to better understand their dynamics. The CCCTC-binding factor (CTCF) is an architectural protein that binds DNA and acts as a barrier factor to define TAD boundaries [120]. As such, genome-editing of CTCF binding sites is a straightforward method to manipulate the looping interactions within TADs [98,121–123]. Two adjacent TADs can be merged by disrupting the CTCF sites that separate them, or new loops can be formed by installing new CTCF sites with genome-editing. These experiments provide evidence that the TAD-scale genome organization can be accurately inferred from the occupancy and orientation of CTCF sites (see Section Loop extrusion) [121,124], although the higher-order genome organization into active and inactive compartments depends on different factors [125].

TAD organization regulates gene expression [125–127]. Several studies have shown that deletion of a CTCF site can result in ectopic enhancer-promoter interactions that affect gene expression [122,123,128,129], establishing the CTCF-mediated looping landscape as an input to transcription. Larger-scale editing of TADs to decrease and increase their size have demonstrated that TADs can act as a buffer against the effect of changing linear distance between an enhancer and promoter [130]. CTCF-site deletion followed by FACS and RNA-FISH demonstrated that CTCF sites can reduce variability in cell-to-cell gene expression levels by mediating enhancer-promoter interactions [131]. Now, pooled CRISPR-Cas9 screening can be employed to edit thousands of sites [96,97,132–134]. By targeting CTCF sites and TAD boundaries, experimentalists can efficiently test the effects of chromatin architecture on gene expression at many more sites across the genome.

The tools of synthetic biology have also been applied to 'force' enhancer-promoter looping without relying on the endogenous CTCF-cohesin looping factors [135,136]. By fusing a domain of the nuclear factor Ldb1 (which forms a complex at the globin LCR enhancer) to a designer zinc finger that binds the B-globin promoter, the authors created a synthetic protein that effectively forces an enhancer-promoter loop and activates transcription. They measured the resultant transcriptional bursts with single molecule RNA-FISH and found the forced loop increased the frequency of transcriptional bursts, but not their size. Intriguingly, recent work in Drosophila showed that an enhancer can drive transcriptional bursts from two distal genes at the same time [137], challenging the traditional cartoon of enhancer-promoter looping. These data suggest the formation of 3D dynamic clusters containing multiple enhancers and promoters. It remains to be seen if this is a general mode of regulation in mammalian cells. In this direction, recent demonstrations of dCas9-mediated forced looping [138] lay the groundwork for higher-throughput efforts to measure the consequences of looping at any location in the genome.

Localization of genes within the nucleus

Measurements

Compartmentalization of the genome into different types of chromatin domains and the spatial positioning of genes within the nucleus are ways of regulating gene expression through localization [139]. Measurements of position effects on gene expression have been made with the TRIP (Thousands of Reporter Integrated in Parallel) method. Thousands of reporter transgenes were randomly integrated into the genome of mESCs [140]. The expression levels of each reporter were linked with the genomic and epigenomic features of their integration site, by sequencing both the genomic DNA and the reporter mRNA. Reporters integrated in lamina-associated domains (LADs) produced 5- to 6-fold less mRNA expression compared to inter-LAD regions. This reduced activity of reporters in LADs may be caused by low density of functional enhancers and reduced binding of transcription factors at LADs.

The molecular mechanisms of gene localization are not yet fully understood. Another method called high-throughput imaging positioning mapping (HIPMap) utilizes DNA-FISH and image analysis to map the 3D position of multiple endogenous loci within the nucleus [141]. Applying HIPMap in a siRNA screen for genome organization factors revealed 50 genes (ranging from chromatin remodelers to nuclear envelope and pore proteins) that are required for positioning, reflecting the complex control of this input to gene expression. Coupling HIPMap with RNA-FISH could provide a complete map of the transcriptional capacity of any position in the nucleus.

In addition to the static methods above, the CRISPR-Cas9 system has been optimized to allow for the dynamic visualization of endogenous loci [142,143]. Using EGFP-tagged dCas9 targeted to specific repetitive genomic loci, these studies monitored telomere dynamics [142,143], as well as the localization of the Mucin genes locus during mitosis [142]. Live tracking of genomic loci via fluorescently-tagged dCas9 coupled to single-cell monitoring of transcription would shed light on the connection between gene localization and expression dynamics.

Manipulations

Synthetic protein fusions enable researchers to forcibly relocate genes to the nuclear periphery for silencing. A specific gene can be relocated by tagging it with a LacO array to recruit an engineered lac repressor (LacI) fused to an inner nuclear membrane protein. This manipulation directly demonstrated transcriptional repression of genes located near the nuclear periphery [144,145]. More recently, two nucleoporin sub-complexes Nup133 and Nup50 were individually fused to LacI-CFP, such that they bound an intranuclear LacO array. This protein fusion recruited additional nucleoporin sub-complexes to the LacO site and, as before, translocated this site to the nuclear periphery [146]. Interestingly, these studies reveal that nucleoporin domains are sufficient to promote gene relocation to the nuclear periphery as compared to full-length nuclear membrane proteins. The combination of programmable DNA binding domains like dCas9 with small localization domains from nucleoporins could enable the targeted relocation of any gene, enabling genome-wide studies of gene localization's contribution to expression dynamics.

Sequence-specific DNA binding proteins also play a role in controlling the spatial positioning of genes. The DNA binding sites of the transcription factor YY1 (Ying-Yang1)

are sufficient for targeting ectopic sites to the nuclear periphery [147]. Relocation of the binding site is reversed upon shRNA-mediated knockdown of YY1 suggesting that YY1 expression is necessary for relocation to occur. Moreover, tethering of the YY1 transcription factor to a reporter locus using the LacI/lacO system caused local accumulation of repressive histone modifications and the relocation of the locus to the nuclear periphery. These results reveal that a diverse input of TFs, repressive chromatin modifications, and lamina proteins are required for repression associated with re-positioning of endogenous genes to the nuclear periphery.

Additionally, in mammalian cells, the heat shock protein 70 (HSP70) loci can move from their location at the nuclear periphery to interchromatin regions of the nucleoplasm, known as nuclear speckles, upon heat shock [148]. Contact of the HSP70 loci to nuclear speckles leads to increased gene expression of these loci. Inserting the HSP70 loci near endogenous genes (by CRISPR for example) would allow for heat-shock inducible control of these genes via change in localization.

Model global input–output function of transcription

For each of the inputs that modulate transcription, there is currently at least one model that should be incorporated into an integrated gene regulation model: the bursting model for connecting transcription factors and chromatin to transcription (Fig. 2a), the chromatin spreading model for relating feedback between chromatin modifications and transcription (Fig. 2b), and the loop extrusion model for describing the dynamical organization of chromatin and its connection with chromatin spreading and enhancer activation (Fig. 2c).

The difficulty with integrating all these models into one coherent global model is that they vary widely in flavor and scope. Some models are molecular – based on physical and biological principles. Molecular models usually start out as cartoons, then advance to dynamic simulations and, rarely, to analytical solutions. The chromatin spreading and loop extrusion models fall in this category. Other models are phenomenological – they are based on fits to experimental data and used to extract key features. The transcriptional bursting model and bioinformatics analysis such as ChromHMM [149] fall in this category. These latter models are very useful to get a handle on the data and to disentangle molecular details in early stages of studying a system, when the molecular details are incompletely understood. They are also very useful approximations to employ on a daily basis, even once the molecular details are clear. Here, we review the current state of these isolated models and make suggestions as to how they could be connected.

Extended transcriptional bursting model

A phenomenological model called the telegraph model was proposed to explain transcriptional bursting (Fig. 2a) [150–152]. Here we describe the two-state telegraph model, and recent developments that extend this model to include the effects of chromatin and transcription factors on mRNA production.

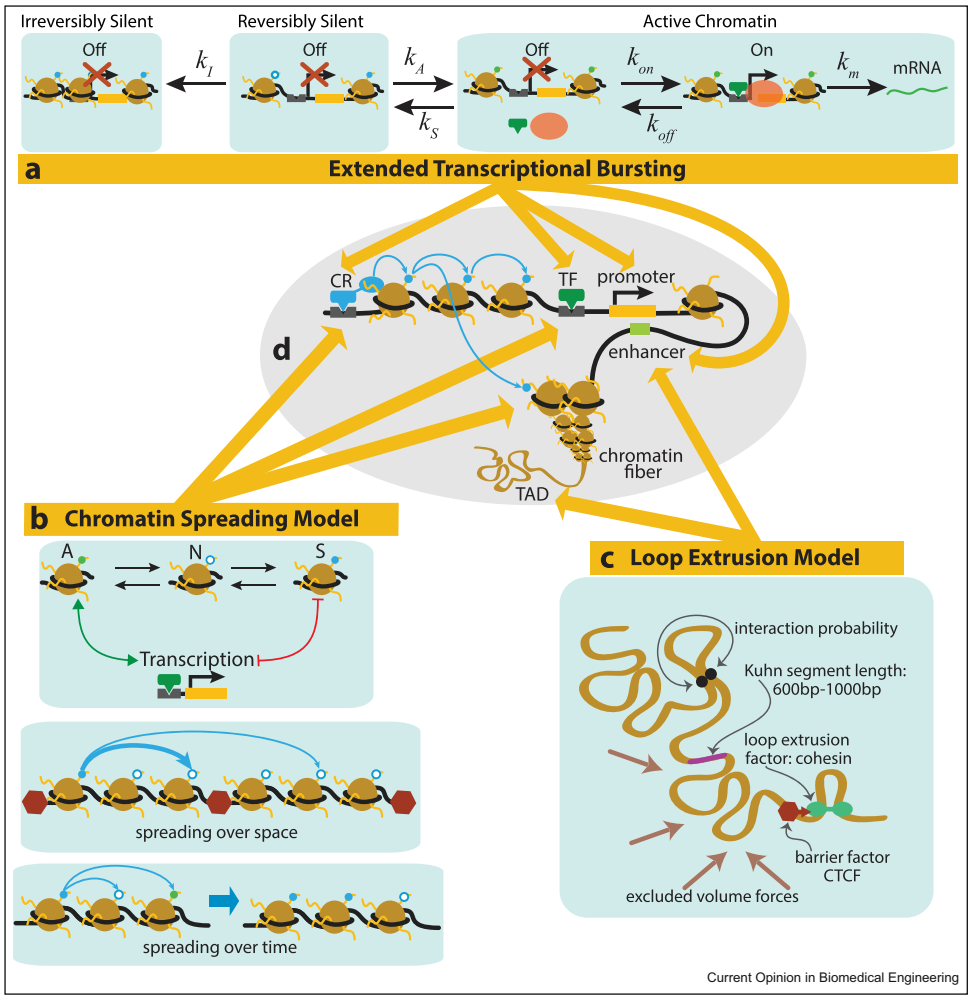
In the original two-state telegraph model, a promoter can switch between two states (ON and OFF) at rates k_{on} and k_{off} (Fig. 2a, active chromatin). Transcription can only initiate from the ON state, and leads to mRNA production at a rate k_m. This effective rate k_m absorbs in it all the rates associated with transcription initiation, elongation, and termination. In this model, the burst frequency is determined by k_{on}, the burst duration is determined by k_{off}, and the burst amplitude by k_m. Hence the burst size (duration times amplitude) depends both on k_{off} and k_m. This model describes transcriptional bursting from a gene located at an active chromatin locus.

The chromatin state of the gene can also undergo stochastic transitions among three states: active, reversibly silent and irreversibly silent (Fig. 2a) [23]. The promoter is OFF in the reversibly and irreversibly silent states, and can switch between ON and OFF in the active state, as described above. The rates of reversible silencing (k_S), reactivation (k_A), and irreversible commitment (k_I) determine the percentage of cells having the gene of interest in an active chromatin state as a function of time, and set the timescale of epigenetic memory at the population level. The rates in this model can be extracted from single-cell measurements of gene expression over time.

Now the questions are: how do the different transcriptional inputs modulate the rates in this model and how do we describe these effects mathematically?

Transcription factors binding at the promoter have been shown to affect the rates related to bursting in the chromatin active state (reviewed in Ref. [1]). Increasing the concentration of an activating TF increases k_{on}, stronger DNA binding domains decrease k_{off}, and stronger activation domains increase k_m [32]. Enhancers mainly control k_{on}, presumably by increasing the local concentration of TFs via looping [136,137]. One simplifying approach to incorporating the effects of TFs into the rate kinetics formalism is to assume that TF binding and enhancer looping are much faster than chromatin and promoter transition rates, and thus can be assumed to reach equilibrium between the steps in the extended telegraph model. In this case, one can mathematically describe the effect of a TF by multiplying the rate it affects by the probability of finding the TF at that locus. However, the experimentally measured response in gene expression to changes in TF concentration is

Fig. 2



Models of gene regulation and chromatin dynamics are connected. (a) Extended transcriptional bursting model (adapted from Ref. [23]). The chromatin switches between active, reversibly silent, and irreversibly silent states. mRNA can only be expressed in the active state. **(b)** Minimal chromatin spreading model that includes transcription (based on [155,157,158]). A nucleosome switches between active, neutral, and silent states. A nucleosome's state affects the probability of its neighboring nucleosomes switching their states. The extent of its influence on neighboring nucleosomes' states may be affected by their 3D interactions and intervening barrier factors. **(c)** The loop extrusion model. A molecular dynamics simulation of the chromatin fiber (brown) folding under excluded volume forces, with additional dynamic loops formed by extrusion and barrier factors, recapitulates features of chromosome conformation maps (adapted from Refs. [121,124]). **(d)** The three main models in (a–c) can be integrated with connections at several elements of chromatin and transcription.

much sharper than would be predicted by theoretical modeling using equilibrium statistics, suggesting that this equilibrium assumption might not be correct [153]. Moreover, while most efforts have concentrated on TF binding in the accessible chromatin state, it is likely that some transcription factors, especially pioneer factors that bind to chromatin, can modulate the transitions between different chromatin states (k_A, k_S, k_I).

Chromatin modifications have been shown to affect all the rates in the extended telegraph model. In the active state, higher acetylation is correlated with increased TF binding kinetics (hence increased k_{on}), as well as increased transcription elongation rates (hence increased k_m) [22]. Complete removal of histone acetylation leads to reversible silencing (k_S) [23]. DNA and repressive histone methylation are associated with the transitions to reversible and irreversible silencing such that writing of DNA methylation leads to irreversible silencing (k_I) [23,81,93] while histone methylation (H3K9me3 and H3K27me3) can affect all three rates (k_S, k_A and k_I) [23,24]. Therefore, histone acetylation plays a role in setting the level of transcription in the active state, while the interplay among

DNA methylation and histone modifications control the fraction of cells in the active state. However, it is not clear how to mathematically connect the microscopic rates of histone and DNA modifications to the phenomenologically measured rates in the kinetic model presented here.

Chromatin spreading

As outlined in the previous sections, chromatin modifications have a strong effect on gene expression. One key feature associated with many types of modifications is their ability to spread across large distances [154]. Multiple groups have modeled this spreading to match the experimentally observed chromatin domain sizes [155–157], and more recently to connect the dynamics of chromatin modification spreading to gene expression [158]. In order to get heritable chromatin domains, a model generally needs two ingredients: 1) having more than two states for each nucleosome (e.g. activating, neutral, and silencing) and 2) positive feedback loops (Fig. 2b). The positive feedback loops consist of modifications recruiting chromatin regulator complexes that write the same modification or remove the opposing modification at other nucleosomes. There are two types of positive feedback: linear spreading, where immediate neighbors can affect each other [156,158], and looping-mediated spreading, where complexes bound at a certain location can contact and modify nucleosomes that are far in linear genomic space [155,157]. While both models can produce stable chromatin domains, only the looping model matches the kinetics of chromatin domains that result from recruitment and removal of chromatin regulators in yeast [157]. Moreover, one can get the appearance of a domain that would suggest spreading even in the absence of positive feedback, simply by tethering a chromatin regulator to DNA [157]. It is therefore important to understand the probability of contact between any two points as a function of distance in order to connect spreading of chromatin modifications to gene expression changes.

Loop extrusion

Measurements of chromatin contacts using Hi-C technology have shown that the contact probability, I, follows a power law dependence on genomic distance, s, such that $I(s) = s^{-\gamma}$ with a scaling coefficient $\gamma = 0.75$ [121]. This phenomenological function can be reproduced by molecular simulations of polymer folding [121,124] using a loop extrusion model. In this molecular model, the chromatin fiber is modeled as a polymer made of monomers each 600–1000 bp. This polymer folds due to excluded volume forces. The model also includes loop extrusion factors (such as cohesin) that can bind the chromatin fiber and produce dynamic loops, and barrier factors (CTCF) that stop the translocation of these loops [121,124]. These simulations recapitulate Hi-C maps and often predict the results of CTCF site editing experiments, in which barrier removal alters the TAD boundaries.

There are still remaining questions in this model about the behavior of chromatin at length-scales smaller than 1000 bp and larger than 10,000,000 bp (nuclear compartments) [124–126] and the role of chromatin modifications and transcription in the contact probability function. Nevertheless, contact probability as a function of genomic distance can be incorporated into the bursting model to describe the dynamics of enhancer looping, and into the chromatin spreading model to calculate the probability of chromatin spreading among non-adjacent points.

Bioinformatics algorithms and functional genomics

Large scale epigenomic mapping efforts, like the Roadmap Epigenomics Consortium and ENCODE, have contributed a wealth of data from diverse cell types and underscore the need for models that can aggregate and interpret diverse data. ChromHMM is one such model; it is a hidden Markov model that learns complex patterns from chromatin modification (obtained from ChIP-seq and other experiments) and then reports interpretable chromatin states (i.e. active promoter, strong enhancer, insulator, etc.) [149,159]. Currently, these methods have been used for the annotation of bulk cell populations, but could be applied to the emerging single-cell epigenomic data [160]. Bioinformatics algorithms are necessary to tease out coarse-grained trends from very large data sets that would impossible to process otherwise. These trends can be used to generate mechanistic hypotheses, and ultimately the results from testing these hypotheses need to be incorporated back in the bioinformatics algorithms.

Conclusions

There has been an explosion of single-cell methods for measuring mammalian gene expression and its determinants – polymerases, TFs, and chromatin state (modifications, accessibility, and 3D architecture). These methods were possible because of advances in microscopy, sequencing, and synthetic biology. In microscopy, super-resolution methods, brighter dyes, and tagging mRNA and proteins with multiple genetically-encoded fluorophores allow for single-molecule detection and fast dynamic measurements (reviewed in Refs. [25,27]). In addition, incubation chambers and reliable auto-focus control have become more common for fluorescent microscopes allowing the imaging of mammalian cells over many days [161]. Next generation sequencing coupled with microfluidics and barcoding methods have enabled single-cell genome-wide measurements of RNA expression [162], chromatin modifications [66], and chromatin accessibility [102,103], albeit with sparser data than from a single locus with microscopy. Up to date, most studies focused on

measuring a single variable at a time; however, both microscopy and sequencing are moving towards measuring both the transcriptional output and its inputs in the same cell.

Advances in mammalian synthetic biology in the last few years include: improved DNA synthesis and delivery, improved genome and epigenome editing with programmable DNA binding domains and protein engineering to enable new drug- or light-inducible functions [26]. These advances have contributed to improving microscopy (by allowing fluorescent tagging of any endogenous protein [163]) and sequencing-based methods (employing synthesis and delivery of large libraries [53]). We are witnessing a systematic, high-throughput, and targeted approach for tuning each of the factors that control transcription.

Despite these great technological advances, we are still far from having a global, coherent, and predictive model of the gene regulatory function. There has been progress modeling isolated aspects of this function at different length scales (i.e. promoters and enhancers for transcriptional bursting, tens to hundreds of nucleosomes for chromatin spreading, and hundreds of kilobases for loop extrusion) and times scales (i.e. seconds to minutes for TF and RNAP binding, and hours to days for chromatin-associated processes). It is necessary to take a systems-wide view and integrate these models across time and space. While post-transcriptional regulation was beyond the scope of this article, its integration will be critical to link transcriptional outputs to phenotypes. Additionally, we need to take into consideration that transcription itself is not only an output, but can modulate inputs by removing histones from DNA, changing supercoiling and chromatin organization, and recruiting additional input factors via the nascent RNA (reviewed in Refs. [164,165]). As such, a complete model of gene regulation will necessarily integrate mechanisms beyond what was discussed here. This global model will provide a better foundation for building gene regulation tools to be used routinely and reliably in basic biology studies and cellular engineering applications.

Acknowledgments

This work was supported by DARPA (W911NF-11-2-0055 – MBE), NIH (NIH R01 HD075335 – MBE; T32 Training Grant T32GM007276 – MVV), NSF (Graduate Research Fellowship – JT), Burroughs Wellcome Fund (Career at the Scientific Interface Award – LB), and HHMI (MBE).

Conflict of interest

None declared.

Glossary of abbreviations

5mc	5-Methylcytosine
5hmc	5-Hydroxymethycytosine
ChIP-Seq	Chromatin Immunoprecipitation with Sequencing
CTCF	CCCTC-binding factor
CRs	Chromatin Regulators
DBDs	DNA Binding Domains
FISH	Fluorescent In Situ Hybridization
NES	Nuclear Export Signal
NLS	Nuclear Localization Signal
LAD	Lamina Associated Domain
PALM	Photoactivated Localization Microscopy
RNAP	RNA Polymerase
sc	Single Cell
STORM	Stochastic Optical Reconstruction Microscopy
TAD	Topologically Associated Domain
TFs	Transcription Factors

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Nicolas D, Phillips NE, Naef F: **What shapes eukaryotic transcriptional bursting?** *Mol Biosyst* 2017, **13**:1280–1290.
2. Sanchez A, Golding I: **Genetic determinants and cellular constraints in noisy gene expression**. *Science* 2013, **342**: 1188–1193 (80-).
3. Lenstra TL, Rodriguez J, Chen H, Larson DR: **Transcription dynamics in living cells**. *Annu Rev Biophys* 2016, **45**:25–47.
4. Vera M, Biswas J, Senecal A, Singer RH, Park HY: **Single-cell and single-molecule analysis of gene expression regulation**. *Annu Rev Genet* 2016, **50**:267–291.
5. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S: **Stochastic mRNA synthesis in mammalian cells**. *PLoS Biol* 2006, **4**:e309.
6. Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L: **Single-cell in situ RNA profiling by sequential hybridization**. *Nat Methods* 2014, **11**:360–361.
7. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X: **Spatially resolved, highly multiplexed RNA profiling in single cells**. *Science* 2015, **348** (80-) aaa6090-aaa6090.
8. Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, Terry R, Jeanty SSF, Li C, Amamoto R, *et al.*: **Highly multiplexed subcellular RNA sequencing in situ**. *Science* 2014, **343**:1360–1363.
9. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA: **The technology and biology of single-cell RNA sequencing**. *Mol Cell* 2015, **58**:610–620.
10. Frieda KL, Linton JM, Hormoz S, Choi J, Chow K-HK, Singer ZS, Budde MW, Elowitz MB, Cai L: **Synthetic recording and in situ readout of lineage information in single cells**. *Nature* 2016, **541**:107–111.
11. Singer ZS, Yong J, Tischler J, Hackett JA, Altinok A, Surani MA, Cai L, Elowitz MB: **Dynamic heterogeneity and DNA methylation in embryonic stem cells**. *Mol Cell* 2014, **55**:319–331.
12. Darzacq X, Shav-Tal Y, de Turris V, Brody Y, Shenoy SM, Phair RD, Singer RH: **In vivo dynamics of RNA polymerase II transcription**. *Nat Struct Mol Biol* 2007, **14**:796–806.
13. Yunger S, Rosenfeld L, Garini Y, Shav-Tal Y: **Single-allele analysis of transcription kinetics in living mammalian cells**. *Nat Methods* 2010, **7**:631–633.
14. Paige JS, Wu KY, Jaffrey SR: **RNA mimics of green fluorescent protein**. *Science* 2011, **333**:642–646.
15. Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA, Yeo GW: **Programmable RNA tracking in live cells with CRISPR/Cas9**. *Cell* 2016, **165**:488–496.
16. Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ, Verdine V, Cox DBT, Kellner MJ, Regev A,

et al.: **RNA targeting with CRISPR–Cas13**. *Nature* 2017, **550**: 280.

17. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F: **Mammalian genes are transcribed with widely different bursting kinetics**. *Science* 2011, **332**:472–474.

18. Sprinzak D, Lakhanpal A, Lebon L, Santat LA, Fontes ME, Anderson GA, Garcia-Ojalvo J, Elowitz MB: **Cis-interactions between Notch and Delta generate mutually exclusive signalling states**. *Nature* 2010, **465**:86–90.

19. Han K, Jaimovich A, Dey G, Ruggero D, Meyuhas O, Sonenberg N, Meyer T: **Parallel measurement of dynamic changes in translation rates in single cells**. *Nat Methods* 2013, **11**:86–93.

20. Cisse II , Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, Dugast-Darzacq C, Hajj B, Dahan M, Darzacq X: **Real-time dynamics of RNA polymerase II clustering in live human cells**. *Science* 2013, **341**:664–667 (80-).

21. * Cho WK, Jayanth N, English BP, Inoue T, Andrews JO, Conway W, Grimm JB, Spille JH, Lavis LD, Lionnet T, *et al.*: **RNA Polymerase II cluster dynamics predict mRNA output in living cells**. *Elife* 2016, **5**:681–686.
Using single-molecule super-resolution microscopy to measure mRNA and Pol II during transcription in live cells, the authors discovered that Pol II forms transient clusters whose lifetimes correlate with the mRNA output.

22. ** Stasevich TJ, Hayashi-Takanaka Y, Sato Y, Maehara K, Ohkawa Y, Sakata-Sogawa K, Tokunaga M, Nagase T, Nozaki N, McNally JG, *et al.*: **Regulation of RNA polymerase II activation by histone acetylation in single living cells**. *Nature* 2014, **516**: 272–275.
The authors used an engineered gene array to simultaneously measure histone modification and polymerase state at the promoter, dynamically, in single-cells. They found that histone acetylation speeds up TF binding kinetics and accelerates polymerase transition to elongation.

23. * Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M: **Elowitz MB: dynamics of epigenetic regulation at the single-cell level**. *Science* 2016, **351**:720–724 (80-).
Here, we showed that recruitment of different chromatin regulators changes gene expression in an all-or-none manner in single cells. We used these data to develop a phenomenological three-state stochastic model of chromatin control.

24. * Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR: **Dynamics and memory of heterochromatin in living cells**. *Cell* 2012, **149**:1447–1460.
The authors measured the temporal dynamics of gene expression and the spreading of chromatin modifications upon recruitment of HP1 at an engineered *Oct4* locus. They used a theoretical model of of chromatin spreading to extract kinetic rates of histone marking and turnover.

25. Cuvier O, Fierz B: **Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells**. *Nat Rev Genet* 2017. https://doi.org/10.1038/nrg.2017.28.

26. Black JB, Perez-Pinera P, Gersbach CA: **Mammalian synthetic biology: engineering biological systems**. *Annu Rev Biomed Eng* 2017, **19**:249–277.

27. Coleman RA, Liu Z, Darzacq X, Tjian R, Singer RH, Lionnet T: **Imaging transcription: past, present, and future**. *Cold Spring Harb Symp Quant Biol* 2015, **80**:1–8.

28. Chen J, Zhang Z, Li L, Chen B-C, Revyakin A, Hajj B, Legant W, Dahan M, Lionnet T, Betzig E, *et al.*: **Single-molecule dynamics of enhanceosome assembly in embryonic stem cells**. *Cell* 2014, **156**:1274–1285.

29. Izeddin I, Recamier V, Bosanac L, Cisse II , Boudarene L, Dugast-Darzacq C, Proux F, Benichou O, Voituriez R, Bensaude O, *et al.*: **Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus**. *Elife* 2014:3.

30. Liu Z, Legant WR, Chen BC, Li L, Grimm JB, Lavis LD, Betzig E, Tjian R: **3D imaging of Sox2 enhancer clusters in embryonic stem cells**. *Elife* 2014, **3**:e04236.

31. Meyer B, Bénichou O, Kafri Y, Voituriez R: **Geometry-induced bursting dynamics in gene expression**. *Biophys J* 2012, **102**: 2186–2191.

32. ** Senecal A, Munsky B, Proux F, Ly N, Braye FEE, Zimmer C, Mueller F, Darzacq X: **Transcription factors modulate c-Fos transcriptional bursts**. *Cell Rep* 2014, **8**:75–83.
They systematically dissected the role of TF properties (concentration, strength of binding and activating domains) on transcriptional bursting parameters extracted from single-molecule RNA-FISH distributions.

33. Suter DM, Molina N, Naef F, Schibler U: **Origins and consequences of transcriptional discontinuity**. *Curr Opin Cell Biol* 2011, **23**:657–662.

34. * Dar RD, Razooky BS, Singh A, Trimeloni TV, McCollum JM, Cox CD, Simpson ML, Weinberger LS: **Transcriptional burst frequency and burst size are equally modulated across the human genome**. *Proc Natl Acad Sci* 2012, **109**:17454–17459.
Quantitative analysis of 8000 loci using time-lapse microscopy revealed that virtually all genes are transcribed by bursting, as opposed to constitutive expression.

35. Rakhit R, Navarro R, Wandless TJ: **Chemical biology strategies for posttranslational control of protein function**. *Chem Biol* 2014, **21**:1238–1252.

36. McKeague M, Wong RS, Smolke CD: **Opportunities in the design and application of RNA for gene expression control**. *Nucleic Acids Res* 2016, **44**:2987–2999.

37. Chung HK, Jacobs CL, Huo Y, Yang J, Krumm SA, Plemper RK, Tsien RY, Lin MZ: **Tunable and reversible drug control of protein production via a self-excising degron**. *Nat Chem Biol* 2015, **11**:713–720.

38. Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AGL, Wandless TJ: **A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules**. *Cell* 2006, **126**:995–1004.

39. Iwamoto M, Björklund T, Lundberg C, Kirik D, Wandless TJ: **A general chemical method to regulate protein stability in the mammalian central nervous system**. *Chem Biol* 2010, **17**: 981–988.

40. Liang F-S, Ho WQ, Crabtree GR: **Engineering the ABA plant stress pathway for regulation of induced proximity**. *Sci Signal* 2011, **4**:rs2.

41. Luo J, Liu Q, Morihiro K, Deiters A: **Small-molecule control of protein function through Staudinger reduction**. *Nat Chem* 2016, **8**:1–13.

42. Robinson MS, Sahlender DA, Foster SD: **Rapid inactivation of proteins by rapamycin-induced rerouting to mitochondria**. *Dev Cell* 2010, **18**:324–331.

43. Niopek D, Benzinger D, Roensch J, Draebing T, Wehler P, Eils R, Di Ventura B: **Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells**. *Nat Commun* 2014, **5**:1–11.

44. Beyer HM, Juillot S, Herbst K, Samodelov SL, Müller K, Schamel WW, Römer W, Schäfer E, Nagy F, Strähle U, *et al.*: **Red light-regulated reversible nuclear localization of proteins in mammalian cells and zebrafish**. *ACS Synth Biol* 2015, **4**: 951–958.

45. Nihongaki Y, Yamamoto S, Kawano F, Suzuki H, Sato M: **CRISPR-Cas9-based photoactivatable transcription system**. *Chem Biol* 2015, **22**:169–174.

46. Polstein LR, Gersbach CA: **Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors**. *J Am Chem Soc* 2012, **134**:16480–16483.

47. Polstein LR, Gersbach CA: **A light-inducible CRISPR-Cas9 system for control of endogenous gene activation**. *Nat Chem Biol* 2015, **11**:198–200.

48. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F: **Optical control of mammalian endogenous transcription and epigenetic states**. *Nature* 2013, **500**:472–476.

49. Kennedy MJ, Hughes RM, Peteya LA, Schwartz JW, Ehlers MD, Tucker CL: **Rapid blue-light-mediated induction of protein interactions in living cells**. *Nat Methods* 2010, **7**:973–975.

50. Renicke C, Schuster D, Usherenko S, Essen LO, Taxis C: **A LOV2 domain-based optogenetic tool to control protein degradation and cellular function**. *Chem Biol* 2013, **20**:619–626.

51. Bonger KM, Rakhit R, Payumo AY, Chen JK, Wandless TJ: **A general method for regulating protein stability with light**. *ACS Chem Biol* 2013, **9**(1):111–115.

52. Stanley SA, Sauer J, Kane RS, Dordick JS, Friedman JM: **Remote regulation of glucose homeostasis in mice using genetically encoded nanoparticles**. *Nat Med* 2015, **21**:92–98.

53. Grossman SR, Zhang X, Wang L, Engreitz J, Melnikov A, Rogov P, Tewhey R, Isakova A, Deplancke B, Bernstein BE, *et al.*: **Systematic dissection of genomic features determining transcription factor binding and enhancer function**. *Proc Natl Acad Sci* 2017. https://doi.org/10.1073/pnas.1621150114.

54. Maricque BB, Dougherty JD, Cohen BA: **A genome-integrated massively parallel reporter assay reveals DNA sequence determinants of *cis*-regulatory activity in neural cells**. *Nucleic Acids Res* 2016, **45**:gkw942.

55. Park M, Keung AJ, Khalil AS: **The epigenome: the next substrate for engineering**. *Genome Biol* 2016, **17**:183.

56. Vora S, Tuttle M, Cheng J, Church G: **Next stop for the CRISPR revolution: RNA-guided epigenetic regulators**. *FEBS J* 2016, **283**:3181–3193.

57. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, *et al.*: **Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex**. *Nature* 2014, **517**:583–588.

58. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ, *et al.*: **Highly efficient Cas9-mediated transcriptional programming**. *Nat Methods* 2015, **12**:326–328.

59. Benveniste D, Sonntag H-J, Sanguinetti G, Sproul D: **Transcription factor binding predicts histone modifications in human cell lines**. *Proc Natl Acad Sci* 2014, **111**:13367–13372.

60. Karlic R, Chung H-R, Lasserre J, Vlahovicek K, Vingron M: **Histone modification levels are predictive for gene expression**. *Proc Natl Acad Sci* 2010, **107**:2926–2931.

61. Dong X, Greven MC, Kundaje A, Djebali S, Brown JB, Cheng C, Gingeras TR, Gerstein M, Guigó R, Birney E, *et al.*: **Modeling gene expression using chromatin features in various cellular contexts**. *Genome Biol* 2012, **13**:R53.

62. Allfrey VG, Faulkner R, Mirksy AE: **Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis**. *Proc Natl Acad Sci U S A* 1964, **51**:786–794.

63. Schwartzman O, Tanay A: **Single-cell epigenomics: techniques and emerging applications**. *Nat Rev Genet* 2015, **16**:716–726.

64. Clark SJ, Lee HJ, Smallwood SA, Kelsey G, Reik W: **Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity**. *Genome Biol* 2016, **17**:72.

65. Gomez D, Shankman LS, Nguyen AT, Owens GK: **Detection of histone modifications at specific gene loci in single cells in histological sections**. *Nat Methods* 2013, **10**:171–177.

66. * Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA: **Bernstein BE: single-cell ChIP-seq reveals cell subpopulations defined by chromatin state**. *Nat Biotechnol* 2015, **33**:1165–1172.

The authors developed Drop-ChIP – a combination of microfluidics, DNA-barcoding, and sequencing; which brings the genome-wide measurement of chromatin-binding proteins to the single-cell scale.

67. Shema E, Jones D, Shoresh N, Donohue L, Ram O, Bernstein BE: **Single-molecule decoding of combinatorially modified nucleosomes**. *Science* 2016, **352**:717–721 (80-).

68. Kriaucionis S, Heintz N: **The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain**. *Science* 2009, **324**:929–930 (80-).

69. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, *et al.*: **Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1**. *Science* 2009, **324**:930–935.

70. Wu TP, Wang T, Seetin MG, Lai Y, Zhu S, Lin K, Liu Y, Byrum SD, Mackintosh SG, Zhong M, *et al.*: **DNA methylation on N6-adenine in mammalian embryonic stem cells**. *Nature* 2016, **532**:329–333.

71. Bird A: **DNA methylation patterns and epigenetic memory**. *Genes Dev* 2002, **16**:6–21.

72. Li E, Zhang Y: **DNA methylation in mammals**. *Cold Spring Harb Perspect Biol* 2014, **6**:a019133.

73. Xiaoji Wu YZ: **TET-mediated active DNA demethylation: mechanism, function and beyond**. *Nat Publ Gr* 2017, **18**:517–534.

74. Medvedeva YA, Khamis AM, Kulakovskiy IV, Ba-Alawi W, Bhuyan MSI, Kawaji H, Lassmann T, Harbers M, Forrest AR, Bajic VB: **Effects of cytosine methylation on transcription factor binding sites**. *BMC Genomics* 2014, **15**:119.

75. Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D: **Competition between DNA methylation and transcription factors determines binding of NRF1**. *Nature* 2015, **528**:575–579.

76. Bahar Halpern K, Vana T, Walker MD: **Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development**. *J Biol Chem* 2014, **289**(34):23882–23892.

77. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suvà ML, Bernstein BE: **Insulator dysfunction and oncogene activation in IDH mutant gliomas**. *Nature* 2015, **529**:110–114.

78. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, Das PK, Kivioja T, Dave K, Zhong F, *et al.*: **Impact of cytosine methylation on DNA binding specificities of human transcription factors**. *Science* 2017, **356**. eaaj2239.

79. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, Münzel M, Wagner M, Müller M, Khan F, *et al.*: **Dynamic readers for 5-(Hydroxy)methylcytosine and its oxidized derivatives**. *Cell* 2013, **152**:1146–1159.

80. Iurlaro M, Ficz G, Oxley D, Raiber E-A, Bachman M, Booth MJ, Andrews S, Balasubramanian S, Reik W: **A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation**. *Genome Biol* 2013, **14**:R119.

81. * Stelzer Y, Shivalila CS, Soldner F, Markoulaki S, Jaenisch R: **Tracing dynamic changes of DNA methylation at single-cell resolution**. *Cell* 2015, **163**:218–229.

These authors established and used a methylation-sensitive fluorescent reporter to track the DNA methylation changes of known pluripotency enhancers during differentiation and reprogramming.

82. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schönegger A, Klughammer J, Bock C: **Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics**. *Cell Rep* 2015, **10**:1386–1397.

83. Cheow LF, Quake SR, Burkholder WF, Messerschmidt DM: **Multiplexed locus-specific analysis of DNA methylation in single cells**. *Nat Protoc* 2015, **10**:619–631.

84. Szwagierczak A, Brachmann A, Schmidt CS, Bultmann S, Leonhardt H, Spada F: **Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine**. *Nucleic Acids Res* 2011, **39**:5149–5156.

85. Mooijman D, Dey SS, Boisset J-C, Crosetto N, van Oudenaarden A: **Single-cell 5 hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction**. *Nat Biotechnol* 2016, **34**:852–856.

86. Song C, Diao J, Brunger AT, Quake SR: **Simultaneous single-molecule epigenetic imaging of DNA methylation and**

hydroxymethylation. *Proc Natl Acad Sci U S A* 2016, **113**: 4338–4343.

87. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA: **Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers**. *Nat Biotechnol* 2015, **33**:510–517.

88. Cano-Rodriguez D, Rots MG: **Epigenetic editing: on the verge of reprogramming gene expression at will**. *Curr Genet Med Rep* 2016, **4**:170–179.

89. Moosmann P, Georgiev O, Thiesen HJ, Hagmann M, Schaffner W: **Silencing of RNA polymerases II and III-dependent transcription by the KRAB protein domain of KOX1, a Krüppel-type zinc finger factor**. *Biol Chem* 1997, **378**: 669–677.

90. Margolin JF, Friedman JR, Meyer WK, Vissing H, Thiesen HJ, Rauscher FJ: **Krüppel-associated boxes are potent transcriptional repression domains**. *Proc Natl Acad Sci U S A* 1994, **91**:4509–4513.

91. Deuschle U, Meyer WK, Thiesen HJ: **Tetracycline-reversible silencing of eukaryotic promoters**. *Mol Cell Biol* 1995, **15**: 1907–1914.

92. Kungulovski G, Nunna S, Thomas M, Zanger UM, Reinhardt R, Jeltsch A: **Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained**. *Epigenet Chromatin* 2015, **8**:12.

93. Amabile A, Migliara A, Capasso P, Biffi M, Cittaro D, Naldini L, Lombardo A: **Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing**. *Cell* 2016, **167**. 219–232.e14.

94. O'Geen H, Ren C, Nicolet CM, Perez AA, Halmai J, Le VM, Mackay JP, Farnham PJ, Segal DJ: **dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression**. *Nucleic Acids Res* 2017, **45**:9901–9916.

95. Thakore PI, Black JB, Hilton IB, Gersbach CA: **Editing the epigenome: technologies for programmable transcription and epigenetic modulation**. *Nat Methods* 2016, **13**: 127–137.

96. Wright JB, Sanjana NE: **CRISPR screens to discover functional noncoding elements**. *Trends Genet* 2016, **32**:526–529.

97. Sanjana NE, Wright J, Zheng K, Shalem O, Fontanillas P, Joung J, Cheng C, Regev A, Zhang F: **High-resolution interrogation of functional elements in the noncoding genome**. *Science* 2016, **353**:1545–1549.

98. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R: **Editing DNA methylation in the mammalian genome**. *Cell* 2016, **167**. 233–247.e17.

99. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE: **High-resolution mapping and? Characterization of open chromatin across the genome**. *Cell* 2008, **132**:311–322.

100. Song L, Crawford GE: **DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells**. *Cold Spring Harb Protoc* 2010. 2010:pdb.prot5384.

101. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ: **Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position**. *Nat Methods* 2013, **10**: 1213–1218.

102. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ: **Single-cell chromatin accessibility reveals principles of regulatory variation**. *Nature* 2015, **523**:486–490.

103. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J: **Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing**. *Science* 2015:348 (80-).

104. Corces MR, Buenrostro JD, Wu B, Greenside PG, Chan SM, Koenig JL, Snyder MP, Pritchard JK, Kundaje A, Greenleaf WJ, *et al.*: **Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution**. *Nat Genet* 2016, **48**:1193–1203.

105. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L: **Genome architectures revealed by tethered chromosome conformation capture and population-based modeling**. *Nat Biotechnol* 2011, **30**:90–98.

106. Ong C-T, Corces VG: **Enhancer function: new insights into the regulation of tissue-specific gene expression**. *Nat Rev Genet* 2011, **12**:283–293.

107. Dekker J, Mirny L: **The 3D genome as moderator of chromosomal communication**. *Cell* 2016, **164**:1110–1121.

108. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping**. *Cell* 2014, **159**: 1665–1680.

109. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289–293.

110. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, Chang HY: **HiChIP: efficient and sensitive analysis of protein-directed genome architecture**. *Nat Methods* 2016, **13**:919–922.

111. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions**. *Nature* 2012, **485**: 376–380.

112. Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteche CM, Noble WS, Duan Z, Shendure J: **Massively multiplex single-cell Hi-C**. *Nat Methods* 2017, **14**:263–266.

113. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P: **Single-cell Hi-C reveals cell-to-cell variability in chromosome structure**. *Nature* 2013, **502**:59–64.

114. Nagano T, Lubling Y, Várnai C, Dudley C, Leung W, Baran Y, Mendelson Cohen N, Wingett S, Fraser P, Tanay A: **Cell-cycle dynamics of chromosomal organization at single-cell resolution**. *Nature* 2017, **547**:61–67.

115. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu C, Zhuang X: **Super-resolution imaging reveals distinct chromatin folding for different epigenetic states**. *Nature* 2016, **529**:418–422.

116. Beliveau BJ, Boettiger AN, Avendaño MS, Jungmann R, McCole RB, Joyce EF, Kim-Kiselak C, Bantignies F, Fonseka CY, Erceg J, *et al.*: **Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes**. *Nat Commun* 2015, **6**:7147.

117. Beliveau BJ, Joyce EF, Apostolopoulos N, Yilmaz F, Fonseka CY, McCole RB, Chang Y, Li JB, Senaratne TN, Williams BR, *et al.*: **Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes**. *Proc Natl Acad Sci U S A* 2012, **109**:21301–21306.

118. Linhoff M, Garg S, Mandel G: **A high-resolution imaging approach to investigate chromatin architecture in complex tissues**. *Cell* 2015, **163**:246–255.

119. Beagrie R, Pombo A: **Examining topological domain influence on enhancer function**. *Dev Cell* 2016, **39**:523–524.

120. Ong C-T, Corces VG: **CTCF: an architectural protein bridging genome topology and function**. *Nat Publ Gr* 2014:15.

121. Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, *et al.*: **Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes**. *Proc Natl Acad Sci U S A* 2015, **112**:E6456–E6465.

122. Hnisz D, Weintraub AS, Day DS, Valton A-L, Bak RO, Li CH, Goldmann J, Lajoie BR, Fan ZP, Sigova AA, *et al.*: **Activation of proto-oncogenes by disruption of chromosome neighborhoods**. *Science* 2016, **351**:1454–1458 (80-).

123. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin D, Jung I, Wu H, Zhai Y, Tang Y, *et al.*: **CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function**. *Cell* 2015, **162**:900–910.

124. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA: **Formation of chromosomal domains by loop extrusion**. *Cell Rep* 2016, **15**:2038–2049.

125. * Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG: **Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization**. *Cell* 2017, **169**. 930–944.e22.

Acute CTCF degradation demonstrated that CTCF is necessary for chromatin loops and TAD insulation, but not necessary for the larger scale of genome organization in compartments. This paper, and other recent work, suggests different factors must be responsible for compartmentalization.

126. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, Haering C, Mirny L, *et al.*: **Two independent modes of chromatin organization revealed by cohesin removal**. *Nature* 2017. https://doi.org/10.1038/nature24281.

127. Zuniga A, Michos O, Spitz F, Haramis A-PG, Panman L, Galli A, Vintersten K, Klasen C, Mansfield W, Kuc S, *et al.*: **Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression**. *Genes Dev* 2004, **18**:1553–1564.

128. de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, Splinter E, Wijchers PJ, Krijger PHL, de Laat W: **CTCF binding polarity determines chromatin looping**. *Mol Cell* 2015, **60**:676–684.

129. Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D: **CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation**. *Science* 2015, **347**:1017–1021 (80-).

130. Symmons O, Pan L, Remeseiro S, Aktas T, Klein F, Huber W, Spitz F: **The Shh topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances**. *Dev Cell* 2016, **39**:529–543.

131. * Ren G, Jin W, Cui K, Rodrigez J, Hu G, Zhang Z, Larson DR, Zhao K: **CTCF-mediated enhancer-promoter interaction is a critical regulator of cell-to-cell variation of gene expression**. *Mol Cell* 2017, **67**:1049–1058. e6.

The authors combine CTCF site deletion by CRISPR-Cas9 with single-cell expression assays to demonstrate how CTCF can reduce expression noise by mediating enhancer-promoter interactions.

132. Sanjana NE: **Genome-scale CRISPR pooled screens**. *Anal Biochem* 2016. https://doi.org/10.1016/j.ab.2016.05.014.

133. Wang T, Wei JJ, Sabatini DM, Lander ES: **Genetic screens in human cells using the CRISPR-Cas9 system**. *Science* 2014, **343**:80–84.

134. Hartenian E, Doench JG: **Genetic screens and functional genomics using CRISPR/Cas9 technology**. *FEBS J* 2015, **282**:1383–1393.

135. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA: **Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor**. *Cell* 2012, **149**:1233–1244.

136. * Bartman CR, Hsu SC, Hsiung CC-S, Raj A, Blobel GA: **Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping**. *Mol Cell* 2016, **62**:237–247.

The authors forced an enhancer loop with a zinc finger fusion protein to demonstrate that increased enhancer-promoter interaction results in increased transcriptional burst frequency, but not increased burst amplitude.

137. Fukaya T, Lim B, Levine M: **Enhancer control of transcriptional bursting**. *Cell* 2016, **166**:358–368.

138. Morgan SL, Mariano NC, Bermudez A, Arruda NL, Wu F, Luo Y, Shankar G, Jia L, Chen H, Hu J, *et al.*: **Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping**. *Nat Commun* 2017, **8**:1–9.

139. Yáñez-Cuna JO, van Steensel B: **Genome–nuclear lamina interactions: from cell populations to single cells**. *Curr Opin Genet Dev* 2017, **43**:67–72.

140. * Akhtar W, De Jong J, Pindyurin AV, Pagie L, Meuleman W, De Ridder J, Berns A, Wessels LFA, Van Lohuizen M, Van Steensel B: **Chromatin position effects assayed by thousands of reporters integrated in parallel**. *Cell* 2013, **154**:914–927.

In this paper, the authors describe an approach for the monitoring the transcriptional activity of more than 27,000 reporters randomly integrated into mESCs. They show that LADs are negative regulators of transcription by reducing access of transcription factors to binding sites.

141. * Shachar S, Voss T, Pegoraro G, Sciascia N, Misteli T: **Identification of gene positioning factors using high-throughput imaging mapping**. *Cell* 2015, **162**:911–923.

The authors developed a high-throughput, automated FISH imaging pipeline for large scale mapping of genomic locations and identified factors required for the positioning of genes though a siRNA screen.

142. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, *et al.*: **Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system**. *Cell* 2013, **155**:1479–1491.

143. Takei Y, Shah S, Harvey S, Qi LS, Cai L: **Multiplexed dynamic imaging of genomic loci by combined CRISPR imaging and DNA sequential FISH**. *Biophys J* 2017, **112**:1773–1776.

144. Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P, Ylstra B, Chubb JR, Bickmore WA: **Recruitment to the nuclear periphery can alter expression of genes in human cells**. *PLoS Genet* 2008:4.

145. Reddy KL, Singh H: **Using molecular tethering to analyze the role of nuclear compartmentalization in the regulation of mammalian gene activity**. *Methods* 2008, **45**:242–251.

146. Schwartz M, Travesa A, Martell SW, Forbes DJ: **Analysis of the initiation of nuclear pore assembly by ectopically targeting nucleoporins to chromatin**. *Nucleus* 2015, **6**:40–54.

147. Harr JC, Luperchio TR, Wong X, Cohen E, Wheelan SJ, Reddy KL: **Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins**. *J Cell Biol* 2015, **208**:33–52.

148. Khanna N, Hu Y, Belmont AS: **Report HSP70 transgene directed motion to nuclear speckles facilitates heat shock activation**. *Curr Biol* 2014, **24**:1138–1144.

149. Ernst J, Kellis M: **ChromHMM: automating chromatin-state discovery and characterization**. *Nat Methods* 2012, **9**:215–216.

150. Ko MSH: **A stochastic model for gene induction**. *J Theor Biol* 1991, **153**:181–194.

151. Peccoud J, Ycart B: **Markovian modeling of gene-product synthesis**. *Theor Popul Biol* 1995, **48**:222–234.

152. Larson DR, Singer RH, Zenklusen D: **A single molecule view of gene expression**. *Trends Cell Biol* 2009, **19**:630–637.

153. * Estrada J, Wong F, DePace A, Gunawardena J: **Information integration and energy expenditure in gene regulation**. *Cell* 2016, **166**:234–244.

Developed a mathematical formalism for connecting TF binding kinetics to transcription. They used this model to show that the high sharpness in gene expression vs. TF concentration observed in eukaryotes can only come about from very strong constraints on TF cooperativity or additional energy expenditure.

154. Talbert PB, Henikoff S: **Spreading of silent chromatin: inaction at a distance**. *Nat Rev Genet* 2006, **7**:793–803.

155. Dodd IB, Micheelsen M a, Sneppen K, Thon G: **Theoretical analysis of epigenetic cell memory by nucleosome modification**. *Cell* 2007, **129**:813–822.

156. Hodges C, Crabtree GR: **Dynamics of inherently bounded histone modification domains**. *Proc Natl Acad Sci U S A* 2012:2012.

157. * Erdel F, Greene EC: **Generalized nucleation and looping model for epigenetic memory of histone modifications**. *Proc Natl Acad Sci U S A* 2016, **113**:E4180–E4189.
The authors simulated linear and looping-driven spreading of chromatin modifications, and showed that the size of engineered chromatin domains in yeast is better aligned with a looping model.

158. * Berry S, Dean C, Howard M: **Slow chromatin dynamics allow polycomb target genes to filter fluctuations in transcription factor activity**. *Cell Syst* 2017, **4**. 445–457.e8.
Developed a theoretical model that includes the effect of transcription on histone turnover and epigenetic memory. They showed that this model can lead to bistability in gene expression without histone modifications spreading via long-range looping.

159. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, *et al.*: **Integrative analysis of 111 reference human epigenomes**. *Nature* 2015, **518**:317–330.

160. Dekker J, Belmont AS, Guttman M, Leshyk VO, Lis JT, Lomvardas S, Mirny LA, O'Shea CC, Park PJ, Ren B, *et al.*: **The 4D nucleome project**. *Nature* 2017, **549**:219–226.

161. Skylaki S, Hilsenbeck O, Schroeder T: **Challenges in long-term imaging and quantification of single-cell dynamics**. *Nat Biotechnol* 2016, **34**:1137–1144.

162. Liu S, Trapnell C: **Single-cell transcriptome sequencing: recent advances and remaining challenges**. *F1000 Research* 2016:5.

163. Roberts B, Haupt A, Tucker A, Grancharova T, Arakaki J, Fuqua MA, Nelson A, Hookway C, Ludmann SA, Mueller IM, *et al.*: **Systematic gene tagging using CRISPR/Cas9 in human stem cells to illuminate cell organization**. *Mol Biol Cell* 2017, **28**:2854–2874. https://doi.org/10.1101/123042.

164. Lai WKM, Pugh BF: **Understanding nucleosome dynamics and their links to gene expression and DNA replication**. *Nat Rev Mol Cell Biol* 2017, **18**:548–562.

165. Skalska L, Beltran-Nebot M, Ule J, Jenner RG: **Regulatory feedback from nascent RNA to chromatin and transcription**. *Nat Rev Mol Cell Biol* 2017, **18**:331–337.

166. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM: **Localization of ASH1 mRNA particles in living yeast**. *Mol Cell* 1998, **2**:437–445.

167. Shav-Tal Y, Singer RH, Darzacq X: **Innovation: imaging gene expression in single living cells**. *Nat Rev Mol Cell Biol* 2004, **5**: 855–862.

168. Kanda T, Sullivan KF, Wahl GM: **Histone GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells**. *Curr Biol* 1998, **8**:377–385.

169. Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Ferrante TC, Terry R, Turczyk BM, Yang JL, Lee HS, Aach J, *et al.*: **Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues**. *Nat Protoc* 2015, **10**: 442–458.

170. Fudenberg G, Imakaev M: **FISH-ing for captured contacts: towards reconciling FISH and 3C Authors**. *Nat Methods* 2017. https://doi.org/10.1101/081448.

171. Cheow LF, Courtois ET, Tan Y, Viswanathan R, Xing Q, Tan RZ, Tan DSW, Robson P, Loh Y, Quake SR, *et al.*: **Single-cell multimodal profiling reveals cellular epigenetic heterogeneity**. *Nat Methods* 2016, **13**:833–836.

172. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, *et al.*: **CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes**. *Cell* 2013, **154**:442–451.

173. Perez-Pinera P, Ousterout DG, Brunger JM, Farin AM, Glass KA, Guilak F, Crawford GE, Hartemink AJ, Gersbach CA: **Synergistic and tunable human gene activation by combinations of synthetic transcription factors**. *Nat Methods* 2013, **10**: 239–242.

174. * Braun SMG, Kirkland JG, Chory EJ, Husmann D, Calarco JP, Crabtree GR: **Rapid and reversible epigenome editing by endogenous chromatin regulators**. *Nat Commun* 2017, **8**:560.
The authors developed an inducible dCas9-based epigenome editing tool to recruit endogenous chromatin regulators (BAF and Hp1/Suv39h1 complexes) to control bivalent gene expression in mESCs.

175. Vojta A, Dobrinic P, Tadic V, Bockor L, Korac P, Julg B, Klasic M, Zoldos V: **Repurposing the CRISPR-Cas9 system for targeted DNA methylation**. *Nucleic Acids Res* 2016, **44**:5615–5628.

176. Cong L, Zhou R, Kuo Y, Cunniff M, Zhang F: **Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains**. *Nat Commun* 2012, **3**:968.

177. Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK, Bernstein BE: **Locus-specific editing of histone modifications at endogenous enhancers**. *Nat Biotechnol* 2013, **31**:1133–1136.

178. Bernstein DL, Le Lay JE, Ruano EG, Kaestner KH: **TALE-mediated epigenetic suppression of CDKN2A increases replication in human fibroblasts**. *J Clin Invest* 2015, **125**: 1998–2006.

179. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, *et al.*: **Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins**. *Nat Biotechnol* 2013, **31**:1137–1142.

180. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P: **Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription**. *Online* 2011, **29**: 149–154.

181. Maeder ML, Linder SJ, Reyon D, Angstman JF, Fu Y, Sander JD, Joung JK: **Robust, synergistic regulation of human gene expression using TALE activators**. *Nat Methods* 2013, **10**: 243–245.