Published online 26 October 2020

Nucleic Acids Research, 2020, Vol. 48, No. 22 ***e127***
doi: 10.1093/nar/gkaa906

TriTag: an integrative tool to correlate chromatin dynamics and gene expression in living cells

Haiyue Xu[1], Junyan Wang[1], Ying Liang[1], Yujuan Fu[1], Sihui Li[1], Jinghan Huang[2], Heng Xu[2,3], Wei Zou[4,5] and Baohui Chen[1,6,7,8,*]

[1]Department of Cell Biology, and Bone Marrow Transplantation Center of the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China, [2]School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China, [3]Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China, [4]The Fourth Affiliated Hospital, Zhejiang University School of Medicine, Yiwu 322000, China, [5]Insititute of Translational Medicine, Zhejiang University, Hangzhou 310003, China, [6]Institute of Hematology, Zhejiang University & Zhejiang Engineering Laboratory for Stem Cell and Immunotherapy, Hangzhou 310058, China, [7]Zhejiang Laboratory for Systems & Precision Medicine, Zhejiang University Medical Center, Hangzhou 311121, China and [8]Zhejiang Provincial Key Laboratory of Genetic & Developmental Disorders, Hangzhou 310058, China

Received March 23, 2020; Revised August 29, 2020; Editorial Decision September 29, 2020; Accepted October 21, 2020

ABSTRACT

A wealth of single-cell imaging studies have contributed novel insights into chromatin organization and gene regulation. However, a comprehensive understanding of spatiotemporal gene regulation requires developing tools to combine multiple monitoring systems in a single study. Here, we report a versatile tag, termed TriTag, which integrates the functional capabilities of CRISPR-Tag (DNA labeling), MS2 aptamer (RNA imaging) and fluorescent protein (protein tracking). Using this tag, we correlate changes in chromatin dynamics with the progression of endogenous gene expression, by recording both transcriptional bursting and protein production. This strategy allows precise measurements of gene expression at single-allele resolution across the cell cycle or in response to stress. TriTag enables capturing an integrated picture of gene expression, thus providing a powerful tool to study transcriptional heterogeneity and regulation.

INTRODUCTION

The establishment of complex transcriptional programs with temporal and spatial precision is crucial for many biological processes (1,2). However, our understanding of how specific genes are spatiotemporally regulated remains incomplete. It is becoming increasingly apparent that chromatin remodeling, transcription regulation, and protein expression are interconnected at the cellular level, thus, it is important to understand how they are coordinated in living cells (3–8).

Specific labeling of biomolecules with fluorescent tags in living cells is the cornerstone of cell biology studies. Fluorescent tools are leading a revolution in illuminating various cellular molecules, including genomic DNA, RNA and protein (9–15). We had previously developed the dCas9-FP system for imaging endogenous genomic DNA, which allows tracking of chromatin dynamics (16,17). The use of the MS2-MCP approach to monitoring multiple steps of the RNA life cycle in real-time has provided unprecedented insights into the process of gene expression (18–21). Analysis of protein turnover by fluorescent probes was widely performed to elucidate gene expression as well (22,23). However, to explore chromatin dynamics and their relation to transcriptional status, it is highly desirable to combine multiple monitoring systems in a single cell.

A pioneering study created a tetracycline-inducible system to monitor gene expression at the levels of DNA, RNA, and protein in living cells (24). This system is composed of a 200-copy transgene array by combining the lac operator/lac repressor to label DNA (31,32), the MS2-MCP system to tag RNA, and the fluorescent tag to report protein. However, every copy of the transgene in this system is ∼20 kb in length, which does not aim at endogenous gene tagging. Here, we report a small tag (∼1.5 kb) which can be genetically inserted into the endogenous locus at the N- or C-terminus of a protein-coding gene by genome editing. This tag enables simultaneous real-time imaging of chromatin dynamics (DNA), transcriptional status (RNA) and protein turnover in a single living cell.

*To whom correspondence should be addressed. Tel: +86 571 88208304; Fax: +86 571 88208022; Email: baohuichen@zju.edu.cn

MATERIALS AND METHODS

Cell culture

HeLa cells and HEK293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) with high glucose (Gibco) in 10% FBS (Hyclone) and 1% penicillin/streptomycin (Gibco). U2OS was cultured in McCoy's 5A (Procell) supplemented with 10% FBS, 1% penicillin and streptomycin. All cells were cultured at 37°C and 5% CO_2 in a humidified incubator. Cells used in this study were maintained in mycoplasma-free status.

Plasmids construction

The Addgene plasmids #40649 plasmid (25) and # 104999 (26) were used to express tdMCP-GFP and stdMCP-stdHaloTag, respectively. The construction of dCas9-GFP11$_{14X}$ and GFP1–10 plasmids has been described in our previous study (27). The following plasmids were specifically constructed for this study.

1) Construction of reporter plasmids
   To build stdMCP-tdTomato, the DNA sequence encodes stdMCP was amplified from an Addgene plasmid #104999 (26). Fragments of both stdMCP and tdTomato were cloned into a lentiviral vector phage-ubc (Addgene #40649) using NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs). To build plasmids for live-cell RNA labeling (Supplementary Figure S1), the DNA fragment 12xMS2V5 amplified from Addgene #84561 (28) was integrated into the 5′UTR or 3′UTR of mCherry in our lentiviral backbone vector pHR-SFFV-mCherry (same backbone with Addgene #80409) using T4 DNA ligase (New England Biolabs). To specifically detect nascent RNAs, the fourth intron of human HSPA5 gene was inserted into mCherry- or BFP-coding sequence. A restriction site BstXI was artificially embedded in the intron region for the ease of molecular cloning. MS2$_{nx}$ (PCR template: Addgene #27118) (18), or MS2V5$_{nx}$ was then cloned into the BstXI site using T4 DNA ligase. To build the doxycycline-inducible system (Supplementary Figure S3), mCherry (intron-MS2V5$_{17X}$) was cloned to the lentiviral Tet-On 3G inducible vector using T4 DNA ligase. To observe transcriptional dynamics of CMV and SFFV promoters, TriTag (BFP) was cloned to the corresponding backbone vectors by T4 DNA ligase.
2) Assembly of TriTags
   To achieve simultaneous imaging of DNA, nascent RNA, and protein of a specific gene, we designed three versions of TriTags, which are hybrids of fluorescent protein, CRISPR-Tag (27) and MS2 loops. The hybrid of CRISPR-Tag and MS2 loops was directly synthesized by GenScript and then cloned into the intron region of BFP or mCherry, resulted in the final generation of TriTag (BFP-TriTag or mCherry-TriTag). The DNA sequence of TriTags are shown in Supplementary Tables S1 and S2. For BFP-TriTag, three versions of TriTags were designed, which differ in the sequence of CRISPR-Tag. Six CRISPR targeting sequences from *C. elegans* genome, which were validated by genome editing (29–31), were selected to assemble CRISPR-Tags as we have described previously (27).
3) Construction of donor plasmids
   To tag an endogenous gene with a fluorescent protein or TriTag, we performed CRISPR-mediated homology-directly repair (HDR). Thus, a donor plasmid is required for the knock-in experiment. All the donor plasmids used in this study were constructed using the same cloning method. For example, to construct the donor plasmid for inserting TriTag into the C-terminus of H2B, we amplified three fragments: the left and right homology arm (HA) of H2B amplified from genomic DNA of HeLa cells, and the TriTag (BFP). To generate donor plasmids harboring sgH2B recognition sites, termed double-cut donor plasmid (32), the sgH2B-targeting sequence together with the PAM sequence (GCGAGCGCCAGGTCCCGGCAGGG) was included in the forward primer of left HA and the reverse primer of right HA. Therefore, sgH2B targeting sequence was tagged to the regions flanking the upstream and downstream HA. Then the three fragments were cloned into the same vector using NEBuilder HiFi DNA Assembly Cloning Kit. The DNA sequence of H2B-TriTag donor is shown in Supplementary Table S3.
4) Construction of sgRNA plasmids
   All the sgRNAs used in this study were constructed by modifying the CRISPRainbow vector (Addgene #75398) (33). Mouse U6 promoter, specific spacer sequence, and the optimized sgRNA scaffold (E+F) (16) were cloned into the backbone vector by NEBuilder Assembly Kit, resulted in a sgRNA plasmid as small as ∼2.5 kb. The spacer sequence determining the target sequence can be changed to recognize a new site by the PCR-based QuikChange cloning method. All sgRNAs used in this study are all listed in Supplementary Table S4.

Lentivirus production and generation of clonal cell lines

To produce lentivirus, HEK293T cells were seeded into 12-well plates. After ∼12 h, cells were transiently transfected with 750 ng lentivirus constructs (dCas9-GFP11$_{14X}$, GFP1–10, stdMCP-tdTomato, tdMCP-GFP or stdMCP-stdHaloTag), 705 ng pCMV-dR8.91 and 87 ng PMD2.G, using FuGENE (Promega) following the manufacture's recommended protocol. Virus was harvested 60 h after transfection, centrifuged at 800g for 8 min to collect supernatant, and directly added to cells or frozen at –80°C. HeLa cells were infected with tdMCP-GFP lentivirus for live-cell RNA labeling. HeLa, 293T and U2OS cells were infected with dCas9-GFP11$_{14X}$, GFP1–10, and stdMCP-tdTomato lentiviruses for TriTag (BFP version) labeling. To achieve TriTag (mCherry version) imaging, HeLa cells were infected with dCas9-GFP11$_{14X}$, GFP1–10 and stdMCP-stdHaloTag lentiviruses. To enhance the infection of each lentivirus, polybrene (5 μg/ml) was used for lentivirus infection. HeLa and 293T clonal cell lines that express these components at an optimal level for the best labeling of RNA or DNA were isolated and used for imaging experiments. These clonal cell lines were selected based on the signal-to-noise ratio of

DNA or RNA labeling. However, cell pool was directly used for TriTag imaging in U2OS cells.

Generation of reporter cell lines

To generate the doxycycline-inducible system for tracking the production of protein and nascent RNAs in real time, clonal cell line of tdMCP-GFP was infected by the lentivirus of pTRE3G-mCherry (intron-MS2V5$_{17X}$). To create a stable cell line for monitoring transcription dynamics of CMV promoter, clonal cells of stdMCP-tdTomato were plated to 24-well plates and transduced with 100 ng pPB-CMV-Tritag (BFP) plasmids and 200 ng transposase expression vectors. On the next day, cells were subjected to hygromycin selection (300 μg/ml). After 2-week-long selection period, cells with BFP expression were enriched and used for further analysis. To create stable cell lines for tracking the transcriptional dynamics of SFFV promoter, clonal cells of stdMCP-tdTomato grown in 24-well plates were infected with the pSFFV-TriTag (BFP) lentivirus. To enhance the infection of each lentivirus, polybrene (5 μg/ml) was used for lentivirus infection.

CRISPR-mediated knock-in

HeLa, 293T, or U2OS cells expressing particular reporters were grown in 24-well plates and transiently transfected with 100 ng Cas9 protein-expressing vectors, 400 ng donor plasmids, and 500 ng sgRNA expression vectors. Clonal cells were used for HeLa and 293T, while U2OS pool cells were directly applied for knock-in experiments. FACS selection of knock-in positive cells was performed 3–4 days after transfection. To further increase the HDR efficiency in HeLa cells, we used a G2/M phase synchronizer, nocodazole (MCE, HY-13520), to treat the cells for 20 h after the transfection of Cas9/sgRNA/donor. Knock-in efficiencies were summarized in Supplementary Table S5.

Flow cytometry

Protein expression was analyzed by flow cytometry on BD Fortessa instrument (BD Biosciences). CRISPR knock-in positive cells were analyzed and sorted on MoFlo Astrios EQ (Beckman). Cells were first gated for the intact cell population based on forward scatter versus side scatter plots and then gated for single cells using forward scatter W versus forward scatter H. BFP- or mCherry-positive cells were finally sorted out to perform further validation of CRISPR knock-in.

Drugs used to evaluate the labeling of nascent RNAs

To validate the labeling of nascent RNAs, we took advantage of triptolide (MCE, HY-32735) which inhibits transcription activity of RNA Pol II polymerase (34), and isoginkgetin (MCE, HY-N2117) which is a Pre-mRNA splicing inhibitor (35). Cells with appropriate reporters were seeded into eight-well chambered coverglass (Thermo Fisher Scientific) prior to drug treatments. The next day, cells were recorded using a spinning-disk confocal microscope in the absence or presence of drugs. Final concentration of triptolide and isoginkgetin were 30 and 40 μM, respectively. To induce protein expression with the doxycycline-inducible system, cells were incubated with doxycycline (1 μg/ml, Sigma-Aldrich) during a live-cell recording period of 6 hours.

Experimental procedure to monitor HSPA1A activation

To monitor the dynamic changes of chromatin structure and transcription induced by the stress condition, HSPA1A-TriTag cells were plated onto 8-well chambered coverglass 12h prior to transfection. To label the DNA of HSPA1A loci, 500 ng sgTS1 plasmids were transfected to the cells using FuGene (Promega) following the manufacture's recommended protocol. After 24 h of transfection, cells were incubated in 42°C water bath for different lengths of time (30 min, 1 h, 2 h, 3 h or 4 h) and then transferred to widefield or confocal microscope for live cell imaging. During recording, cells were kept at 42°C in a humidified chamber. However, CO_2 was not additionally provided. Thus, there was ∼0.03% CO_2 in the chamber. To observe the reversible chromatin change, cells were incubated in 42°C water bath for 3 h and were then transferred to the confocal microscope for live cell imaging. Cells were kept at 37°C and 5% CO_2 in the humidified chamber.

Identification of sister chromatids

To define a pair of sister chromatids, we isolated clonal cell lines from the TriTag-LMNA knock-in cell pool. Based on DNA imaging, a clonal cell line (*sc1*) with only one-allele tagged by TriTag was selected for measuring the distance between paired sister chromatids. The distance ranges from 0.46 to 1.16 μm. Besides, another clonal cell line (*sc10*), which harbored two TriTag-modified LMNA alleles, was isolated based on DNA and RNA imaging. We quantified the distance between two alleles which locate on two homologous chromosomes. In this case, their distance ranges from 4.28 up to 19.07 μm. Thus, the distance between sister chromatids is much smaller than that between homologous alleles. Based on these parameters, we could precisely define paired sister chromatids for the quantification in Figure 4D.

Quantitative RT-PCR

To examine whether TriTag insertion affects the transcription of target genes in the TriTag imaging system, corresponding cells were collected using trypsin (Hyclone). Total RNA was then extracted using FastPure Cell/Tissue Total RNA Isolation Kit (Vazyme) following the manufacturer's instructions. RNA was converted to cDNA using oligo-dT primers (HiScript II Q RT SuperMix for qPCR, Vazyme). PCR reactions were prepared using ChamQ Universal SYBR qPCR Master mix (Vazyme) and were performed on the QuantStudio 5 Real-Time PCR system (Thermo Fisher). All reactions were done at least in triplicate. RNA abundance was normalized to an endogenous reference gene UBC and calculated as delta-delta threshold cycle (ΔΔCt). Primers used for H2B and LMNA q-RT-PCR, respectively, were listed in our previous study (26).

Widefield microscopy

All widefield microscopy images were performed on a Nikon Ti2-E fluorescence microscope equipped with a ×100 NA 1.45 PlanApo oil immersion objective, an LED source (SPECTRA 4), an sCMOS camera (ZYLA 4.2MP Plus), a Perfect Focus Unit (Nikon), and a motorized stage (Nikon) with stage incubator (Tokai Hit, STRF-WELSX-SET). Cells were grown in 8-well chambered coverglass for live cell imaging. For live-cell imaging, cells were maintained at 37°C and 5% CO_2 in a humidified chamber. The following images were acquired on this widefield microscope: Figure 3D and Supplementary Figures S6, S7, S19, S21. The following quantifications were analyzed based on widefield images: Figures 2D-E, 3H and Supplementary Figures S1C–D, S6B–C, S7C, S15B–C, S19B, S21B–C, S22D–F.

Confocal microscopy

All confocal images were acquired on an Olympus IX83 fluorescence microscope equipped with spinning-disk confocal scanner (Yokogawa CSU-W1), a ×60 NA 1.49 oil Apochromat objective, an sCMOS camera (Prime 95B), 405/488/561/640 nm lasers (OBIS), and a PEIZO stage (ASI) with stage incubator (Tokai Hit). For live-cell imaging, cells were maintained at 37°C and 5% CO_2 in a humidified chamber. Cells for confocal imaging were plated into 8-well chambered coverglass. The following images were acquired on this confocal microscope: Figures 2B, 3A–B, 4A, D, H, 5A, F, H and Supplementary Figures S1B, S3, S9, S10B, S11, S12, S13, S14A, S16, S18, S22. The following quantifications were analyzed based on confocal images: Figures 2C, F, 3A–C, E–G, 4B–C, E–G, I–J, 5B–E, G, I and Supplementary Figures S3, S9, S10, S11, S12, S13, S14, S15A, S16E–F, S17, S18, S20, S22B–C. All supplementary movies were taken on the spinning-disk confocal microscope. *Z* stack images were processed by projection of maximum intensity to generate the movies.

DNA tracking and analysis

To analyze the mobility of individual LMNA loci with or without transcription activity, high-frequency short-term imaging (0.25 s per frame, 400 frames) data was collected on the spinning-disk confocal microscope. The GFP spots in each frame were determined by fitting the image with a Gaussian function by the Object Tracking module of CellSens software (Olympus). Trajectories were created by linking identified puncta to their nearest neighbors. Only trajectories last for 400 frames were used for mobility analysis. The object tracking data generated by CellSens was imported into MATLAB (The MathWorks, Natick, MA, USA) for calculation. For each trajectory, the mean square displacement (MSD) as a function of time delay $t = n\Delta t$ was calculated by the following formula:

$$\mathrm{MSD}\ (n\Delta t) = \frac{1}{N-1-n}\sum_{i=1}^{N-1-n} |\boldsymbol{r}\,(i\Delta t + n\Delta t) - \boldsymbol{r}\,(i\Delta t)|^2,$$

where Δt is the frame length (0.25 s), n is the number of frames in a time delay, N is the total number of frames and $\boldsymbol{r}(t)$ is the two-dimensional coordinate. We then developed a program written in MATLAB to perform the analysis of MSD curves. The shortest 75 time delays (0–15 s) of individual MSD curves were fitted by least-squares regression to a model for confined diffusion, macroscopic diffusion and active transport:

$$\mathrm{MSD}\ (t) = A\left(1 - \mathrm{e}^{-t/\tau}\right) + 4D_{\mathrm{macro}}t + v^2t^2,$$

where A is the confinement area, τ is a constant from which the microscopic diffusion coefficient $D_{\mathrm{micro}} = A/4\tau$\$can be derived, D_{macro} is the macroscopic diffusion coefficient and v is the velocity of active transport. The curves were fitted with a constraint for positive parameters A, τ, v, and D_{macro}. The MSD curves were averaged for the display in Figure 5C. The methods used here to calculate MSD have been described in our previous work (16).

Data analysis

All the fluorescence imaging data were analyzed by ImageJ to calculate the mean intensity, total intensity and the area of genomic loci. The area of genomic loci was defined by the size of dCas9-GFP_{14X} spots, which was calculated based on the signal intensity over background. GraphPad Prism (Version 5, GraphPad Software, La Jolla, CA, USA, https://www.graphpad.com) was used to calculate the mean/median values, coefficient of determination (R^2), correlation coefficient (r), and the standard error of the mean (SEM) for the statistical analysis. Line scan was performed using the 'Analyze/Plot Profile' function, a plugin for ImageJ. The parameters were then analyzed in Excel and plotted in GraphPad Prism. Signal-to-noise ratio was calculated as the ratio of the intensity of a fluorescent signal and the power of background noise as the following formula:

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}} = \frac{\text{Max intensuty of GFP spot} - \text{Mean intensity of background GFP}}{\text{stdDev of background signal}}$$

RESULTS

Labeling nascent transcripts with the Intron-MS2-MCP system

To label RNA in living cells, MS2 loops were mainly inserted into the untranslated region (UTR) of an mRNA of interest. Therefore, measured mRNAs represent integration over the mRNA lifetime, including nascent transcripts and mature mRNAs, contributed from multiple alleles. Our first goal was to selectively label nascent transcripts near the genomic loci by implementing the MS2-MCP system. To this end, we took advantage of the fact that introns are typically degraded rapidly after being spliced out (36–39). We generated a mCherry reporter, harboring 12 copies of MS2V5 loops in 5′ UTR, 3′ UTR and its intron, respectively (Supplementary Figure S1A). 12xMS2V5 is a non-repetitive version of MS2 loops that showed improvement in reproducibility in single-RNA detection (40). Tandem MS2 coat proteins (tdMCP), which bind MS2 RNA loops, were fused to GFP to generate tdMCP-GFP for RNA labeling (25). As expected, UTR tagging resulted in visible spots of tdMCP-GFP in both nucleus and cytoplasm, which represents both nascent and mature RNAs of mCherry. However, the intron of nascent RNAs coated by tdMCP-GFP

enabled specific detection of nascent transcripts in the nucleus, indicating that the transgene of mCherry reporter was actively transcribing (Supplementary Figure S1B and C). Notably, mCherry expression was significantly reduced with the incorporation of MS2 loops in the 5′ or 3′ UTR region but remained normal if MS2 loops were inserted in the intron (Supplementary Figure S1D). Similar results were also observed using endogenous tagging of human HIST2H2BE (encoding histone H2B) and LMNA genes (Supplementary Figure S2).

Three additional assays were performed to further ascertain that we were indeed imaging nascent transcripts in the nucleus using the Intron-MS2-MCP system. First, stdMCP-tdTomato (the synonymous tdMCP) spots disappeared as short as 8 min following triptolide treatment which inhibits RNA Pol II activity (Supplementary Figure S3A–C). Second, perturbing splicing by isoginkgetin treatment led to a significant accumulation of tdMCP-GFP at transcription sites (Supplementary Figure S3D–F). Finally, we built a doxycycline-inducible system to track the expression dynamics of mCherry, which was engineered with MS2 loops in its intron. Here tdMCP-GFP spots appeared in response to the addition of doxycycline. Remarkably, mCherry protein expression was detected ∼1 h after the first wave of transcriptional burst, revealing the dynamic flow of gene expression from RNA to protein (Supplementary Figure S3G–J and Supplementary Movie S1). Taken together, our results demonstrate that the Intron-MS2-MCP system selectively illuminates the production of nascent transcripts in the nucleus and does not significantly impact protein expression.

Design of TriTags for imaging gene expression in real time

Next, we sought to combine DNA, nascent RNA, and protein imaging systems in a single cell. CRISPR-Tag is a new generation of DNA tagging system which can be efficiently recognized by dCas9-FP coupled with 1–4 highly active sgRNAs (27). We generated a hybrid of CRISPR-Tag and MS2 loops, and further assembled the hybrid tag within the intron region of blue fluorescent protein (BFP). A single tag was therefore created to perform triple tagging of DNA, RNA, and protein via a single experiment. Thus, we name this new tag as TriTag, which is about 1.5 kb in size. In principle, TriTag can be applied to image the dynamic flow of genetic information within individual cells, that is, the cellular process of DNA transcribed into mRNA and mRNA translated into protein (Figures 1 and 2A).

To image DNA, RNA, and protein simultaneously, we first isolated a clonal HeLa cell line which stably expressed an appropriate level of dCas9-GFP_{14X} and stdMCP-tdTomato. Furthermore, we created three versions of TriTags (TriTag 1, 2 and 3), which can be labeled by dCas9-GFP_{14X} with different numbers of sgRNAs (Supplementary Figure S4). To test TriTag labeling, we inserted TriTag into the C-terminus of the human H2B gene by CRISPR–Cas9 assisted homologous recombination. BFP expression was used as a FACS sorting marker to select CRISPR knock-in positive cells (Supplementary Figure S5). TriTag insertion was validated by examining the subcellular localization of H2B-BFP (Figure 2B). All three TriTags include 12 copies of MS2V5 loops, thus enabling comparable labeling efficiency of H2B nascent RNAs (Supplementary Figure S6). However, TriTag 1 appeared to perform best signal-to-noise ratio for DNA labeling by using one efficient sgRNA and was therefore used in the subsequent experiments (Supplementary Figure S7).

Monitoring the TriTag imaging system

To demonstrate the use of TriTag, we examined the labeling of human H2B and LMNA. The genomic loci, irrespective of their transcription status, can be directly visualized via dCas9-GFP_{14X}. Most cells in the population show one GFP spot, representing the number of successfully modified alleles. Transcription sites of both H2B and LMNA genes were visible as stdMCP-tdTomato spots which were co-labeled by dCas9-GFP_{14X} in the nucleus (Figure 2B, C). stdMCP-tdTomato binds to MS2 loops in the intron of nascent RNA but would be quickly removed from mRNA because of splicing. Thus, the intensity of stdMCP-tdTomato spots could reflect the degree of transcriptional activity. Base on this principle, LMNA alleles exhibit higher transcription level than H2B alleles (Figure 2D). Consistent with a previous study (41), the intensity of dCas9-GFP_{14x} spots was significantly decreased during the period when nascent RNAs were actively produced at a high level, especially at LMNA loci (Figure 2E). Our observations suggest that the dCas9 complex likely dynamically associates and disassociates with the target gene perhaps due to the competitive binding of RNA Pol II machinery.

Close examination of the size of dCas9-GFP_{14X} spots, we found that the dCas9-GFP_{14X} spot appeared as a larger structure when LMNA gene was highly active. The average size of dCas9-GFP_{14X} spots, potentially reporting the chromatin architecture of LMNA loci, was increased by a factor of 6 (from ∼0.12 μm^2 for inactive loci to ∼0.76 μm^2 for active sites), which might be indicative of a switch of chromatin status or chromatin environment. However, the size of dCas9-GFP_{14X} spots was not well correlated with transcriptional status at H2B loci (Figure 2F). Notably, the transcriptional level of H2B gene is generally much lower than LMNA gene as revealed by the TriTag system. Thus, dramatic change of dCas9-GFP_{14X} spots in size might be limited to highly active loci, which would be further investigated in subsequent experiments. Furthermore, our results suggest that TriTag tagging strategy had no significant effect on mRNA abundance and protein expression of H2B or LMNA compared to BFP tagging based on quantitative PCR and FACS quantifications (Figure 2G and Supplementary Figure S8).

To further assess the capabilities of TriTag imaging, we monitored seven more protein-coding genes in HeLa cells. Expected subcellular localization of the corresponding protein was observed by fluorescent imaging. DNA, nascent mRNA, and protein could be simultaneously visualized for any of these target genes (Supplementary Figure S9). By measuring fluorescence intensity, we observed that TriTag insertion did not significantly affect protein expression, at least for all the genes that were examined, including LMNA, H2B, HSPA1A, HSPB8, HPDL, and SLC38A2. However, the protein expression of some target genes was dramat-

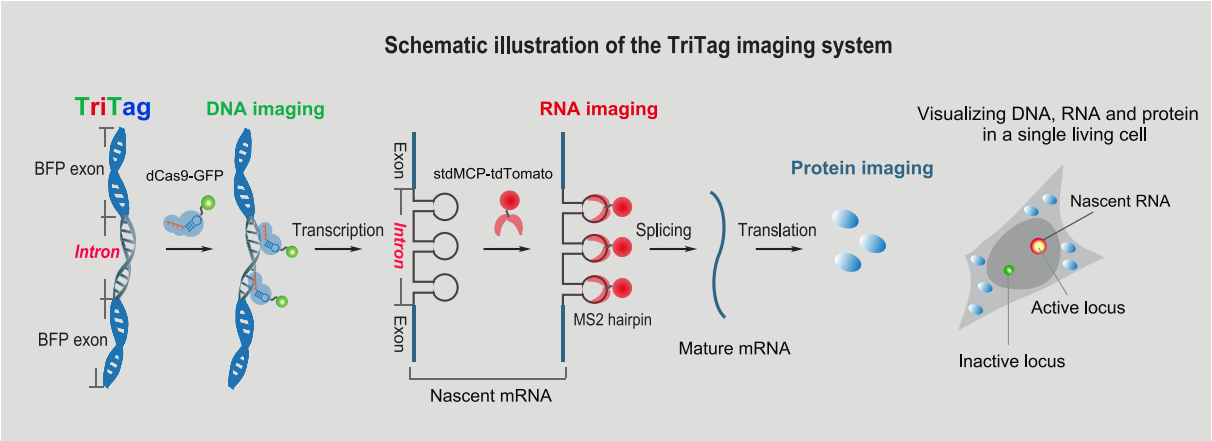

Figure 1. Schematic illustration of the TriTag strategy. Schematic representation of the TriTag imaging system for visualizing DNA, nascent RNA and protein in a single living cell. This system combines three imaging capabilities, including DNA labeling via the CRISPR–Cas9 system, RNA tracking by the MS2-MCP approach, and protein imaging using fluorescent proteins. As an example, a hybrid DNA sequence of CRISPR-Tag and MS2 aptamer is embedded in the intron of blue fluorescent protein (BFP) coding sequence. Transcription is a process where the DNA of a gene is copied to make nascent messenger RNA (mRNA). Nascent mRNA is then spliced to produce mature mRNA, which is later used as instructions to make proteins. CRISPR-Tag is recognized by dCas9-GFP at the DNA level, whereas nascent RNA is bound by stdMCP-tdTomato at the RNA level. Then, the fluorescent protein is expressed. Finally, DNA, nascent mRNA, and protein can be visualized in the same cell by fluorescent imaging.

ically reduced when MS2 loops were incorporated in the UTR region, including 5′ UTR engineering of LMNA and 3′ UTR modifications of H2B, HSPA1A, or HPDL loci (Supplementary Figure S10). These results support the idea that the TriTag system could become a valuable tool for monitoring gene expression.

Additionally, H2B, LMNA, and HSPA1A were also visualized via TriTag in two other human cell lines, HEK293T and U2OS cells (Supplementary Figure S11). It is important to note that TriTag can be assembled as any fluorescent protein with a particular color. Thus, besides BFP-TriTag, we developed the mCherry-TriTag imaging system to further expand the versatility of this approach (Supplementary Figure S12). We next sought to explore whether DNA labeling in the TriTag imaging system affects the transcription of target genes by quantitative PCR. The results revealed no significant difference between samples with and without DNA labeling. Additionally, quantifications based on fluorescent imaging demonstrated that protein expression of target genes was not obviously altered by DNA labeling (Supplementary Figure S13). Together, these results suggest that TriTag strategy is applicable to any protein-coding genes that can be genetically modified by the CRISPR approach in various cell lines. Furthermore, our data demonstrate that TriTag allows precise quantitative measurements of gene expression at the levels of DNA, RNA, and protein, thus providing an integrative view of gene expression.

Connecting transcriptional bursting with protein expression

We next specifically monitored transcriptional bursts of H2B and LMNA. Real-time imaging of stdMCP-tdTomato revealed that the transcription of both genes occurs as a series of bursts, interspaced by silent intervals where no or little mRNA is being produced (Figure 3A and B). Transcriptional bursting has been suggested to be a general phenomenon governing the expression of most eukaryotic genes (42–44). Here, we analyzed a series of burst features in our system, including burst frequency, burst duration and burst amplitude. We found that H2B loci generated ∼2 bursts per hour, whereas the duration of ∼70% of LMNA RNA bursts was longer than 1h. The mean burst duration was 8.4 min for H2B and 125.3 min for LMNA. In addition, the burst amplitude of LMNA is larger than that of H2B (Figure 3C, Supplementary Figure S14 and Supplementary Movie S2). Our quantitative results show that, although both H2B and LMNA are house-keeping genes, they displayed distinct bursting kinetics.

To precisely correlate transcriptional bursting with protein expression, we need to drive the expression of a specific protein using promoters with distinct transcriptional activities. Therefore, we can then measure transcriptional bursting and protein expression for each promoter and correlate these two factors. For this purpose, we selected two promoters, CMV and SFFV, which are broadly used to achieve high expression levels of exogenous proteins. CMV is a much stronger promoter than SFFV in HeLa cells based on our imaging experiments. The bursting analysis revealed that the transcription of both promoters is discontinuous (Figure 3D, E). However, the burst duration of the CMV promoter was much longer than that of SFFV, 24.8 min versus 4.5 min. Additionally, the average pause duration was 2.9 and 8.7 min for CMV and SFFV, respectively. More nascent RNAs were produced by CMV in each burst, elucidated by the intensity of RNA spots (Figure 3F-G and Supplementary Figure S15). We found a tight positive correlation between mRNA production and protein expression (Pearson correlation coefficient, $r = 0.67$ for SFFV, $r = 0.64$ for CMV). The level of BFP (TriTag) expression driven by CMV was indeed much higher than that of SFFV (Figure 3H). These quantitative measurements suggest that the transcriptional bursting is indeed a critical mode of gene regulation, which is also supported by previous studies

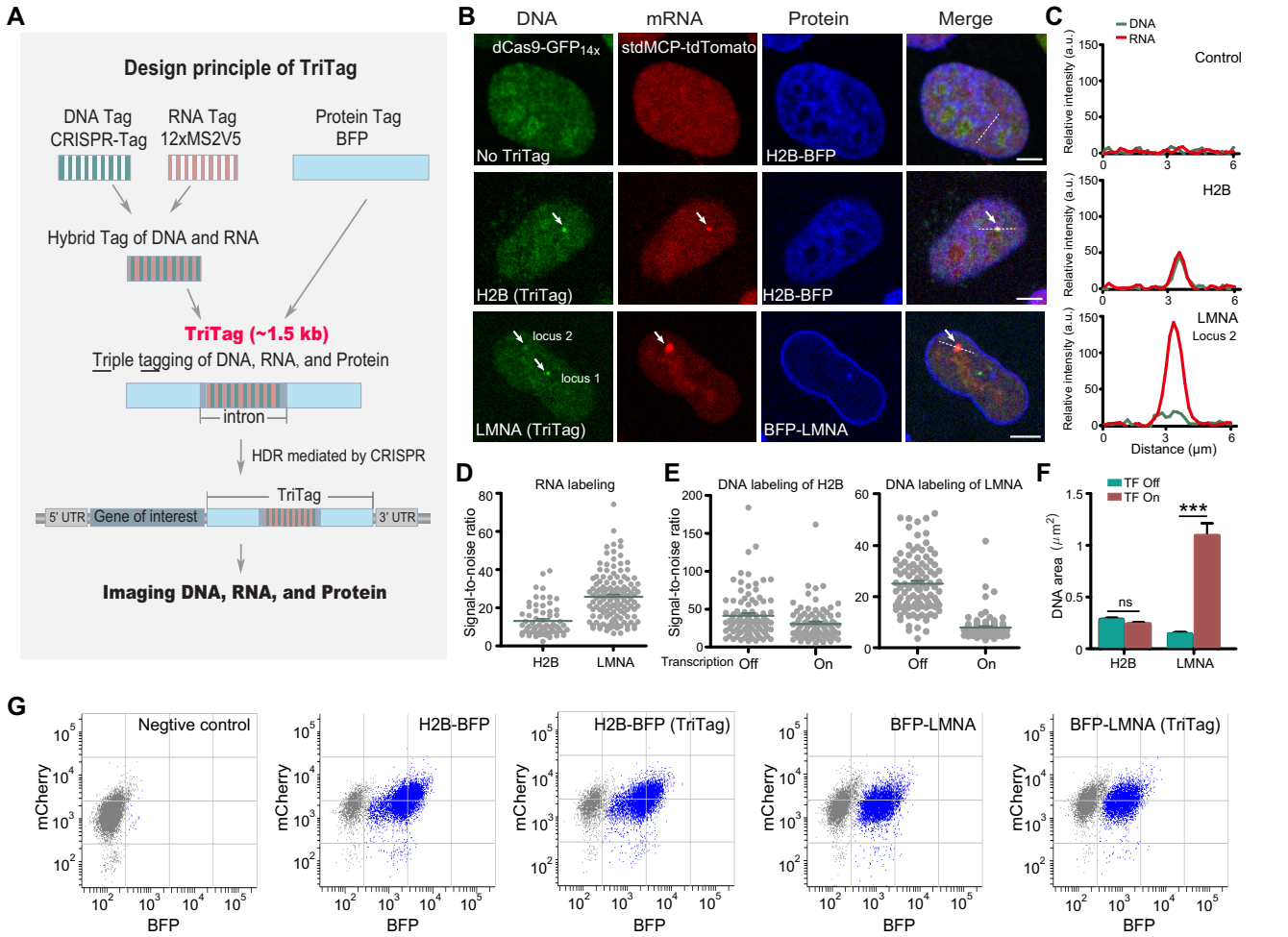

Figure 2. Tricolor visualization of DNA, RNA, and protein in living cells using TriTag. (**A**) Schematic of TriTag design. A hybrid of MS2 loops and CRISPR-Tag was first generated and then inserted in the intron of blue fluorescent protein (BFP). (**B**) Representative images to show simultaneous visualization of DNA, RNA and protein of H2B (top and middle) or LMNA (bottom) genes in HeLa cells. All images are from single focal plane. Scale bars: 5 μm. (**C**) Line scan of the relative fluorescence of the signal indicated by the dotted lines in (C). (**D**) Statistics of signal-to-noise ratio (SNR) to report the labeling of nascent transcripts at H2B and LMNA loci, n ≥ 90 genomic loci. Each dot represents a genomic locus. Green line denotes means ± SEM. (**E**) Statistics of signal-to-noise ratio to demonstrate DNA labeling efficiency at H2B and LMNA loci with or without transcription respectively, $n \geq 97$ genomic loci. Each dot represents a genomic locus. Green line denotes means ± SEM. (**F**) Bar graph showing the size of dCas9-GFP$_{14X}$ spots representing H2B or LMNA loci, with and without transcription, respectively. $n \geq 50$ genomic loci. Error bar represents means ± SEM. ***$P < 0.0001$. (**G**) FACS analysis of H2B-BFP or BFP-LMNA expression in HeLa cells. mCherry serves as an irrelevant channel. The distribution of dots along the *x*-axis indicates BFP signal intensity. Each dot represents a single cell.

(42,45). In summary, our findings demonstrate that TriTag can be applied to quantitatively study transcriptional bursting at endogenous loci in living cells, which could better capture the intrinsic bursting characteristics of eukaryotic genes.

Transcription kinetics across the cell cycle

Since TriTag allows quantifications of allele-specific transcripts, we performed real-time imaging to elucidate how it changes along the cell cycle. To define paired sister chromatids, we measured the distance between paired sister chromatids in a clonal cell line in which only one copy of LMNA gene was tagged with TriTag. In the meanwhile, the distance between two homologous TriTag-LMNA alleles was also calculated in another clonal cell line. These two clonal cell lines were validated by performing PCR experiments and imaging assays, including combinational imaging of DNA, RNA, and protein (Supplementary Figure S16A–E). Our quantifications indicated that the distance between sister chromatids ranges from 0.46 to 1.16 μm, which is much smaller than that between two homologous alleles (Supplementary Figure S16F). Based on these measurements, we identified two dCas9-GFP spots as paired sister chromatids or two homologous alleles for tracking their transcriptional dynamics. First, we specifically imaged transcription kinetics of LMNA gene in those cells harboring two modified homologous alleles in interphase. At a certain time point, there were two distinct states: active or silent expression of both alleles was found in 50% of cells, and monoallelic expression occurred in the other 50% of cells (Figure 4A–C and G). Transcriptional activity mea-

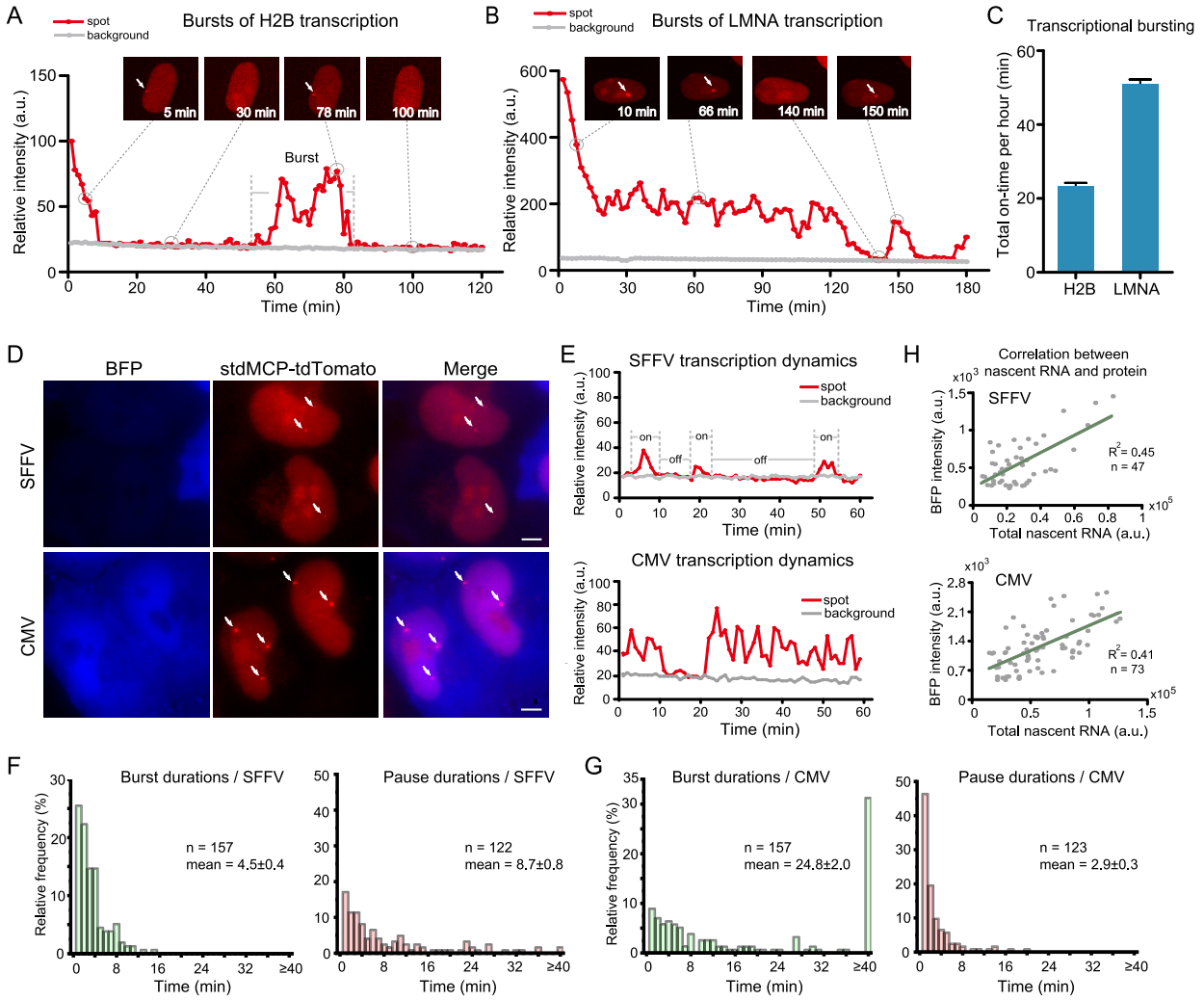

Figure 3. TriTag enables real-time imaging of RNA synthesis and protein expression. (**A**, **B**) Representative traces (red) of nascent transcripts enriched at H2B (A) or LMNA (B) loci compared to that of background signal (gray), illustrating their bursting kinetics. Some snapshots at indicated time points are taken from the cell that was used to generate the corresponding traces. See Movie S2 for dynamics of H2B transcriptional bursting. (**C**) Bar graph showing the average value of total duration time of all bursts occurred within one hour at H2B or LMNA loci, respectively, n ≥ 50 cells. Error bars represent means ± SEM. (**D**) Representative images illustrating the simultaneous labeling of BFP protein and its nascent RNAs. The expression of BFP was driven by two promoters, SFFV (top) and CMV (bottom). All images are maximum intensity projections from *z* stacks. Scale bars: 5 μm. (**E**) Representative traces (red) demonstrating the transcriptional bursting patterns of SFFV and CMV, respectively. Gray traces indicate the intensity of the background signal. (**F**) Histograms of burst durations (n = 157 bursts) and pause durations (*n* = 122 pauses) of SFFV promoter. (**G**) Histograms of burst durations (*n* = 157 bursts) and pause durations (*n* = 123 pauses) of CMV promoter. (**H**) Scatter plots of BFP expression level and the amount of total nascent transcripts in the same cell. Green line denotes the linear fit. R^2 represents the coefficient of determination. Each dot represents a single cell, *n* ≥ 47 cells.

surements indicate a weak anti-correlation between the two alleles within the same cell at a specific time ($r = -0.09$) (Supplementary Figure S17 and Supplementary Movie S3). Next, we sought to examine the transcriptional states of sister chromatids. ∼90% of pairs showed synchronous transcription, both on or both off, suggesting a positive correlation between them ($r = 0.73$). This phenomenon indicates that the 'daughter site' can quickly adopt the transcriptional state of the 'mother site' at the region of LMNA gene (Figure 4D–G).

We next asked whether the transcriptional states can be transmitted from mother cells to daughter cells through mitosis. The transcription kinetics were followed through mitosis captured by time-lapse imaging. We observed that the transcription was shut down at the time of nuclear membrane rupture and recovered in telophase. This phenomenon that mitosis can abort transcription was also reported in previous studies (46,47). Our quantitative analysis revealed a tight correlation between the transcription amplitude of mother cells and daughter cells ($r = 0.77$), suggesting that transcription states of mother cells can be precisely transmitted to daughter cells (Figure 4H–J; Supplementary Movies S4 and S5). This mitotically heritable transcriptional state likely reflects the heritable chromatin states through mitosis. This process might be regulated via epigenetic regulation. Our findings support the idea that chromatin itself maintains critical information during genomic replication (48,49).

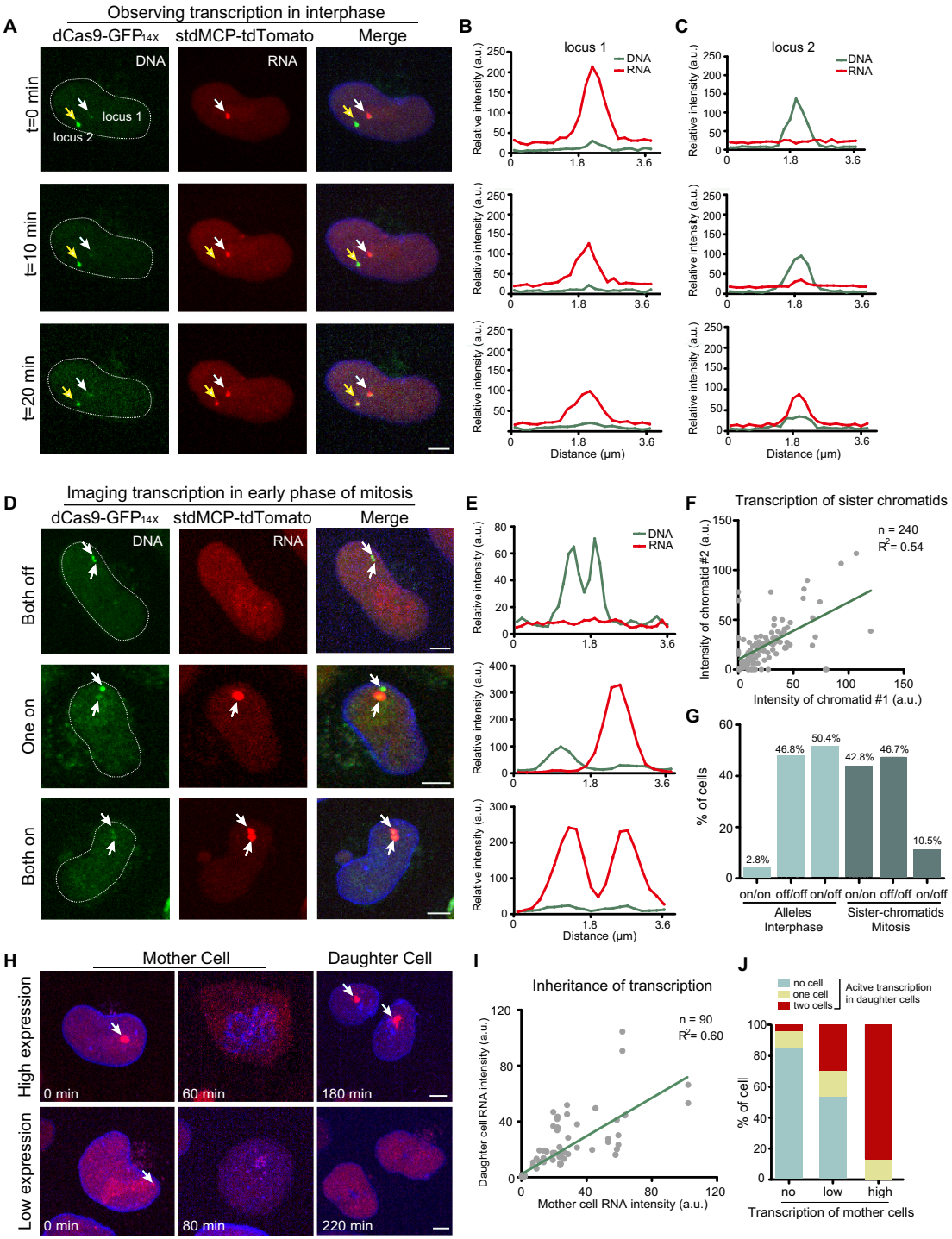

Figure 4. TriTag monitors allele-specific transcriptional bursting across the cell cycle. (**A**) Fluorescent images of a cell showing DNA labeling of two LMNA alleles labeled by dCas9-GFP$_{14X}$ and their transcriptional kinetics indicated by stdMCP-tdTomato at multiple time points. See Movie S3 for dynamics. (**B, C**) Line scan of intensity profiles illustrating dynamic labeling of DNA (green) and nascent RNAs (red) of locus 1 (B) and locus 2 (C) indicated in (A). (**D**) Representative images revealing three distinct transcriptional states of sister chromatid pairs identified by co-labeling of dCas9-GFP$_{14X}$ and stdMCPtdTomato. (**E**) Line scan of intensity profiles showing dynamic labeling of DNA and nascent RNAs at corresponding sister chromatids pointed by arrows in (D). (**F**) Sister chromatids of the same allele have correlated transcription activity. Intensities of both transcription sites were plotted, with each dot representing a different pair of sister chromatids, n = 240 pairs. Green line denotes the linear fit. R^2 represents the coefficient of determination. (**G**) Bar graph showing transcription states of two LMNA alleles in the same cell (case in A, n = 284 cells) or pairs of sister chromatids (case in D, n = 229 cells). (**H**) Snapshots of the transcription activity of LMNA gene in HeLa cells at indicated time points through mitosis. See Movie S4 and S5 for dynamics. (**I**) Transcriptional activity of the mother cell and its daughter cells at LMNA loci was quantified to generate the scatter plot. Each dot represents a pair of the mother cell and one of its daughter cells, $n = 90$ pairs. Green line denotes the linear fit. R^2 represents the coefficient of determination. (**J**) Bar graph showing quantifications of transcriptional activity in daughter cells derived from three groups of mother cells with no, low and high transcription of LMNA gene, respectively. n ≥ 52 cells. All scale bars: 5 μm.

We then quantified the lag period from the first observation of RNA burst in one daughter cell to the time the first burst occurred in the other daughter cell. The time lags in 70% of cases were as short as 20 min. Thus, the timing of first transcriptional bursting in the two daughter cells is relatively close (Supplementary Figure S18). In summary, all the above observations reveal that the TriTag system enables the visualization and single-allele analysis of transcription at every stage in the cell cycle, which would be specifically useful for investigating transcriptional memory (50).

Transcription-coupled chromatin dynamics

We further expanded TriTag to explore the chromatin dynamics and their changes in relation to transcriptional status. High-frequency time-lapse imaging (0.25 s per frame) was performed to monitor the movement of LMNA loci. Trajectory analysis of dCas9-GFP_{14x} spots revealed that LMNA loci displayed confined movement at short time scales (<15 s) (Figure 5A–C). Notably, active loci with abundant transcripts showed significantly lower mobility and smaller confinement area than the inactive loci (Median confinement area: active loci, $D = 0.39 \times 10^{-2}$ μm^2; inactive loci, $D = 1.34 \times 10^{-2}$ μm^2. Corresponding median diffusion coefficient is 1.0×10^{-4} and 2.2×10^{-4} μm^2/min, respectively) (Figure 5D and E). This phenomenon is consistent with previous findings in human and mouse cells (41,51).

We next examined the chromatin dynamics during transcriptional activation. HSPA1A gene, a member of the human heat shock protein 70 family, was tagged with TriTag. Heat-shock treatment resulted in detectable activation of HSPA1A transcription as short as 30 min. The amplitude of transcriptional bursting gradually increased over time under the heat-stress condition (42°C). Strong transcriptional activation was observed between 3 and 4 h (Figure 5F and Supplementary Figure S19A, B). The duration of 'on' state was increased from 7.3 to 53.4 min, whereas the 'off' state duration was reduced from 6.0 min to 2.2 min (Supplementary Figure S20). Next, we measured HSPA1A protein levels after 10 h recovery from heat shock. Our results revealed that protein expression level rose gradually with the increase of exposure time to heat stress, indicating a positive correlation between protein expression and transcriptional level (Supplementary Figure S21A-B). Interestingly, we found that the higher the HSPA1A-BFP protein was expressed, the larger number of HSPA1A-BFP puncta was formed in the cytoplasm (Supplementary Figure S21C). It seems that the formation of HSPA1A puncta was induced by the high concentration of protein.

Interestingly, the size of dCas9-GFP_{14X} spots was obviously enlarged and the mean intensity of the GFP signal was gradually decreased after heat shock. Compare to the normal condition (37°C), the size of dCas9-GFP_{14X} spots was increased by a factor of 5 (from 0.16 to 0.79 μm^2) at 3 h after heat shock. Real-time imaging of the same cell to track both DNA and nascent RNAs revealed that the change of DNA labeling occurred before the transcriptional activation. Changes of DNA labeling were noticeable immediately after heat shock, and fully extended dCas9-GFP_{14X} puncta could be observed at 3–4 h (Figure 5F, G and Supplementary Movie S6). However, cells that were treated with RNA Pol II inhibitor to block the transcriptional activation showed no change in DNA labeling, indicating that the formation of enlarged dCas9-GFP_{14X} puncta is dependent upon ongoing transcription (Figure 5H, I and Supplementary Movie S7). Moreover, dCas9-GFP_{14X} spots could be recovered to the original state after the transcriptional activation was reduced by putting the cells back to the 37°C condition (Supplementary Figure S22A–E). The quantitative analysis implies a significant positive correlation between the size of dCas9-GFP_{14X} spots and the total amount of nascent transcripts indicated by stdMCP-tdTomato ($r = 0.60$) (Supplementary Figure S22F).

According to these findings, together with those observations of active LMNA loci, we propose that the expanded size of dCas9-GFP spots represent a chromatin environment that can facilitate transcriptional activation to the greatest extent. Previous studies reported that doxycycline-induced transcriptional activation was coupled with chromatin decondensation (23,24,52). Thus, accumulating evidence suggests that rapid remodeling of the chromatin structure, which could enhance its accessibility for transcription factors binding, enables robust transcriptional activation. Furthermore, our data also suggest that transcription does seem to play a role in chromatin organization and dynamics. Collectively, our results indicate a direct link between chromatin dynamics and transcriptional regulation, which could be reported by the TriTag imaging system in real time.

DISCUSSION

The human genome encodes over 20,000 protein-coding genes (53). The implementation of fluorescent protein labeling technologies is providing valuable insights into the subcellular localization of proteins and their dynamic behaviors in wide-ranging biological processes (54,55). Recent advances in genome editing medicated by the CRISPR-Cas9 system enable precise genetic modifications, which greatly enhance the fluorescent protein tagging at endogenous loci (56–58). A growing number of applications have demonstrated the power of endogenous gene tagging with the availability of Cas9 endonuclease (59–62). In our study, we have created a new tag, which could in principle be used as a fluorescent protein tag. However, in addition to specific protein labeling, TriTag permits imaging of its nascent transcripts and genomic locus simultaneously in the same cell. The relatively small size of this tag makes it possible for feasible uses in live-cell imaging studies.

An ideal approach to investigate gene regulation is to see its dynamic nature in a living context. Furthermore, studying spatiotemporal control of gene expression accompanied by quantitative analysis of chromatin dynamics would bring a new level of understanding of the underlying mechanisms. Despite the remarkable advances of nucleic acid (both DNA and RNA) live imaging tools over the past decade, relatively few studies have demonstrated simultaneous monitoring of transcription and chromatin dynamics and were mainly limited to study transgenes (24,63,64). Our proof-of-principle study indicates that the TriTag system allows imaging chromatin dynamics at endogenous gene loci with active or inac-

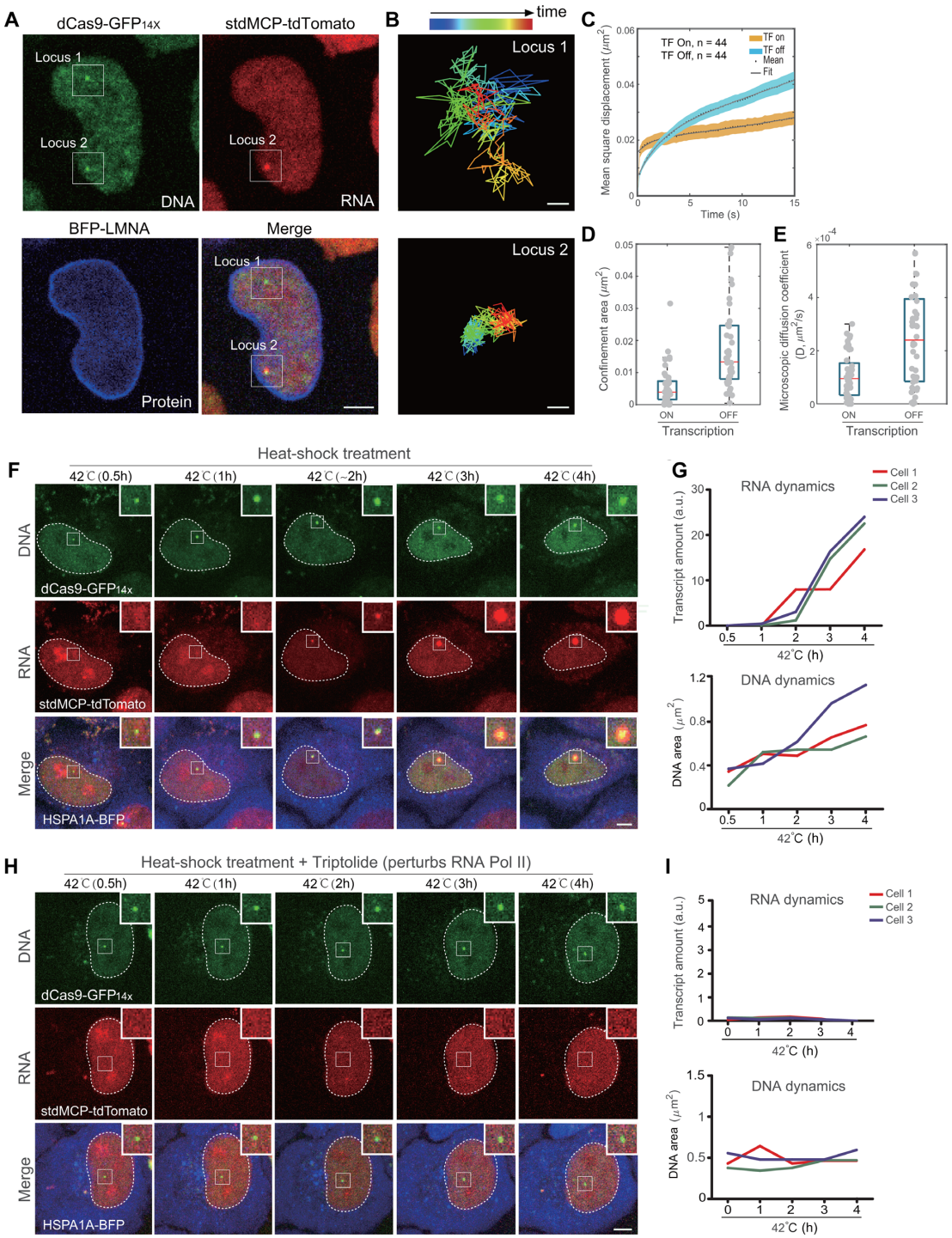

Figure 5. Transcription-coupled chromatin dynamics revealed by the TriTag system. (**A**) Colocalization of dCas9-GFP$_{14X}$ and stdMCP-tdTomato demonstrates the transcription status of LMNA alleles. BFP-LMNA was captured to highlight the nuclear membrane. Scale bar: 5 μm. (**B**) Movement trajectories of dCas9-GFP$_{14X}$ spots representing the two indicated loci in (A). The trajectory lengths are 400 frames with 0.25 s per frame. Scale bars: 100 nm. (**C**) The averaged mean square displacement (MSD) curves of LMNA loci with (TF on) or without (TF off) transcription. The data are displayed as means ± SEM for 44 pairs of alleles. Each pair was from the same cell, with one active and the other one silent. The shaded areas represent the SEM. (**D, E**) Quantifications of the confinement area (D) and the microscopic diffusion coefficient (E) estimated from the MSD curves in (C). Red line at the center of each box denotes the median value, top and bottom edges indicate the 25th and 75th percentiles, respectively. (**F**) Live-cell imaging snapshots of a cell showing chromatin dynamics and transcriptional activation of the HSPA1A gene under the stress condition (42°C, 0.03% CO_2). See Movie S6 for dynamics. Scale bar: 5 μm. (**G**) Quantifications of the transcription activity (top) defined by the total intensity of stdMCP-tdTomato spots, and chromatin dynamics (bottom) by measuring the size of dCas9-GFP$_{14X}$ spots, respectively. Three representative cells were measured, including the one from (F). (**H**) Live-cell imaging snapshots of a cell illustrating chromatin dynamics and transcription activation with the addition of triptolide under the same stress condition as (F). See Movie S7 for dynamics. Scale bar: 5 μm. (**I**) Quantifications were performed as those in (G) for the condition in (H). Three representative cells were analyzed, including the one from (H).

tive transcription, providing robust and quantitative information on gene regulation with spatial and temporal resolution. Tagging more protein-coding genes with TriTags would be the next step to get a better picture of how gene expression is dynamically regulated.

The application of our TriTag imaging system is currently limited to protein-coding genes that can be tagged with TriTag. The knock-in efficiency might not be high. However, the positive cells could be efficiently selected by FACS according to the fluorescent marker. To improve CRISPR-mediated HDR efficiency, we can apply small molecule compounds, such as Nocodazole (65) which worked well in our system. In addition, the types of HDR donors, the length of the homologous arm (32,66), and the TriTag design can be further optimized. Based on DNA imaging, we found that most of the knock-in positive cells only got one allele modified by TriTag. This is likely due to the relatively low knock-in efficiency. Clonal cell lines can be isolated from the pool of knock-in positive cells to get homogeneous genotypes. Next, the number of correctly modified alleles can be validated by imaging assays and quantitative PCR based experiments (67). Random integration of TriTag, which may occur as a rare event, should be considered when examining the genotypes. However, on-target monoallelic editing can be easily confirmed by imaging protein subcellular localization of the target gene, which is a unique advantage of the TriTag imaging system.

Of note, it seems that our TriTag strategy has no significant effect on transcription, mRNA splicing, and protein expression of the target genes that we have tested in this study. To further confirm these characteristics, a systematic examination of more protein-coding genes needs to be conducted. Although the MS2-MCP approach has been extensively applied to monitor the transcription activity in living cells, MS2 loops were mainly inserted into the UTR region of an mRNA of interest. We now embed MS2 loops in the intron of TriTag, which allows specific imaging of nascent RNAs at dCas9-GFP labeled genomic loci. Therefore, our strategy detects nascent RNAs more precisely. Through developing the TriTag imaging system in three human cell lines, we noticed that isolating clonal cell lines to achieve optimal expression level of DNA and RNA labeling components is a critical step to get the best signal-to-noise for imaging. In the RNA imaging system, diffusive MCP-FP proteins, which do not bind to nascent RNAs, act as background signal. Thus, it is likely that the accumulation of nascent RNAs could be efficiently detected only when the gene is being transcribed at a relatively high level. Our observations suggest that there is still room for improvement in the detection sensitivity of transcriptional dynamics.

The advances in the fluorescent tagging toolbox, toward integrating every aspect of the central dogma (DNA to RNA to protein) in a single study, would open a door to address a wide range of biological questions related to gene regulation, such as imprinting genes, transcriptional memory, and transcription-replication conflicts (50,68,69). With the development of CRISPR techniques to edit mammalian genomes, all tools are now available to examine spatiotemporal gene expression at endogenous loci within the context of living cells.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Prof. Jun Ma for critical comments on the manuscript, members of Chen and Zou laboratories for fruitful discussions, Core Facilities of Zhejiang University School of Medicine, especially the Fluorescence Activated Cell Sorting (FACS) center, for the technical support. We gratefully acknowledge the computing resources provided by the student innovation center at Shanghai Jiao Tong University.

FUNDING

National Natural Science Foundation of China [31872819 to B.C., 11774225 to H.X.]; Zhejiang Natural Science Fund for Distinguished Young Scholars [LR20C070003 to B.C.]; National Science Foundation of Shanghai [18ZR1419800 to H.X.]. Funding for open access charge: National Natural Science Foundation of China [31872819].
Conflict of interest statement. None declared.

REFERENCES

1. Levine,M. and Davidson,E.H. (2005) Gene regulatory networks for development. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 4936–4942.
2. Raj,A. and van Oudenaarden,A. (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, **135**, 216–226.
3. Voss,T.C. and Hager,G.L. (2014) Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.*, **15**, 69–81.
4. Xu,H., Sepúlveda,L.A., Figard,L., Sokac,A.M. and Golding,I. (2015) Combining protein and mRNA quantification to decipher transcriptional regulation. *Nat. Methods*, **12**, 739–742.
5. Dekker,J., Belmont,A.S., Guttman,M., Leshyk,V.O., Lis,J.T., Lomvardas,S., Mirny,L.A., O'Shea,C.C., Park,P.J., Ren,B. *et al.* (2017) The 4D nucleome project. *Nature*, **549**, 219–226.
6. Stadhouders,R., Filion,G.J. and Graf,T. (2019) Transcription factors and 3D genome conformation in cell-fate decisions. *Nature*, **569**, 345–354.
7. van Steensel,B. and Furlong,E.E.M. (2019) The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.*, **20**, 327–337.
8. Mateo,L.J., Murphy,S.E., Hafner,A., Cinquini,I.S., Walker,C.A. and Boettiger,A.N. (2019) Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature*, **568**, 49–54.
9. Lippincott-Schwartz,J. and Patterson,G.H. (2003) Development and use of fluorescent protein markers in living cells. *Science*, **300**, 87–91.
10. Buxbaum,A.R., Haimovich,G. and Singer,R.H. (2015) In the right place at the right time: visualizing and understanding mRNA localization. *Nat. Rev. Mol. Cell Biol.*, **16**, 95–109.
11. Chen,B., Guan,J. and Huang,B. (2016) Imaging specific genomic DNA in living cells. *Annu. Rev. Biophys.*, **45**, 1–23.
12. Knight,S.C., Tjian,R. and Doudna,J.A. (2018) Genomes in Focus: development and applications of CRISPR-Cas9 imaging technologies. *Angew. Chem. Int. Ed. Engl.*, **57**, 4329–4337.
13. Wu,X., Mao,S., Ying,Y., Krueger,C.J. and Chen,A.K. (2019) Progress and challenges for Live-cell imaging of genomic loci using CRISPR-based platforms. *Genomics Proteomics Bioinformatics*, **17**, 119–128.
14. Sato,H., Das,S., Singer,R.H. and Vera,M. (2020) Imaging of DNA and RNA in living eukaryotic cells to reveal spatiotemporal dynamics of gene expression. *Annu. Rev. Biochem.*, **89**, 159–187.
15. Shaban,H.A. and Seeber,A. (2020) Monitoring the spatio-temporal organization and dynamics of the genome. *Nucleic Acids Res.*, **48**, 3423–3434.

16. Chen,B., Gilbert,L.A., Cimini,B.A., Schnitzbauer,J., Zhang,W., Li,G.W., Park,J., Blackburn,E.H., Weissman,J.S., Qi,L.S. *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, **155**, 1479–1491.
17. Chen,B., Hu,J., Almeida,R., Liu,H., Balakrishnan,S., Covill-Cooke,C., Lim,W.A. and Huang,B. (2016) Expanding the CRISPR imaging toolset with Staphylococcus aureus Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.*, **44**, e75.
18. Bertrand,E., Chartrand,P., Schaefer,M., Shenoy,S.M., Singer,R.H. and Long,R.M. (1998) Localization of ASH1 mRNA particles in living yeast. *Mol. Cell*, **2**, 437–445.
19. Larson,D.R., Zenklusen,D., Wu,B., Chao,J.A. and Singer,R.H. (2011) Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*, **332**, 475–478.
20. Skinner,S.O., Xu,H., Nagarkar-Jaiswal,S., Freire,P.R., Zwaka,T.P. and Golding,I. (2016) Single-cell analysis of transcription kinetics across the cell cycle. *Elife*, **5**, e12175.
21. Tutucci,E., Livingston,N.M., Singer,R.H. and Wu,B. (2018) Imaging mRNA in vivo, from birth to death. *Annu. Rev. Biophys.*, **47**, 85–106.
22. Giepmans,B.N., Adams,S.R., Ellisman,M.H. and Tsien,R.Y. (2006) The fluorescent toolbox for assessing protein location and function. *Science*, **312**, 217–224.
23. Tsukamoto,T., Hashiguchi,N., Janicki,S.M., Tumbar,T., Belmont,A.S. and Spector,D.L. (2000) Visualization of gene activity in living cells. *Nat. Cell Biol.*, **2**, 871–878.
24. Janicki,S.M., Tsukamoto,T., Salghetti,S.E., Tansey,W.P., Sachidanandam,R., Prasanth,K.V., Ried,T., Shav-Tal,Y., Bertrand,E., Singer,R.H. *et al.* (2004) From silencing to gene expression: real-time analysis in single cells. *Cell*, **116**, 683–698.
25. Wu,B., Chao,J.A. and Singer,R.H. (2012) Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys. J.*, **102**, 2936–2944.
26. Voigt,F., Zhang,H., Cui,X.A., Triebold,D., Liu,A.X., Eglinger,J., Lee,E.S., Chao,J.A. and Palazzo,A.F. (2017) Single-Molecule quantification of translation-dependent association of mRNAs with the endoplasmic reticulum. *Cell Rep.*, **21**, 3740–3753.
27. Chen,B., Zou,W., Xu,H., Liang,Y. and Huang,B. (2018) Efficient labeling and imaging of protein-coding genes in living cells using CRISPR-Tag. *Nat. Commun.*, **9**, 5065.
28. Wu,B., Eliscovich,C., Yoon,Y.J. and Singer,R.H. (2016) Translation dynamics of single mRNAs in live cells and neurons. *Science*, **352**, 1430–1435.
29. Kim,H., Ishidate,T., Ghanta,K.S., Seth,M., Conte,D. Jr, Shirayama,M. and Mello,C.C. (2014) A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. *Genetics*, **197**, 1069–1080.
30. El Mouridi,S., Lecroisey,C., Tardy,P., Mercier,M., Leclercq-Blondel,A., Zariohi,N. and Boulin,T. (2017) Reliable CRISPR/Cas9 genome engineering in Caenorhabditis elegans using a single efficient sgRNA and an easily recognizable phenotype. *G3 (Bethesda)*, **7**, 1429–1437.
31. Farboud,B. and Meyer,B.J. (2015) Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics*, **199**, 959–971.
32. Zhang,J.P., Li,X.L., Li,G.H., Chen,W., Arakaki,C., Botimer,G.D., Baylink,D., Zhang,L., Wen,W., Fu,Y.W. *et al.* (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.*, **18**, 35.
33. Ma,H., Tu,L.C., Naseri,A., Huisman,M., Zhang,S., Grunwald,D. and Pederson,T. (2016) Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.*, **34**, 528–530.
34. Vispe,S., DeVries,L., Créancier,L., Besse,J., Bréand,S., Hobson,D.J., Svejstrup,J.Q., Annereau,J.P., Cussac,D., Dumontet,C. *et al.* (2009) Triptolide is an inhibitor of RNA polymerase I and II-dependent transcription leading predominantly to down-regulation of short-lived mRNA. *Mol. Cancer Ther.*, **8**, 2780–2790.
35. O'Brien,K., Matlin,A.J., Lowell,A.M. and Moore,M.J. (2008) The biflavonoid isoginkgetin is a general inhibitor of Pre-mRNA splicing. *J. Biol. Chem.*, **283**, 33147–33154.
36. Martin,R.M., Rino,J., Carvalho,C., Kirchhausen,T. and Carmo-Fonseca,M. (2013) Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. *Cell Rep.*, **4**, 1144–1155.
37. Levesque,M.J. and Raj,A. (2013) Single-chromosome transcriptional profiling reveals chromosomal gene expression regulation. *Nat. Methods*, **10**, 246–248.
38. Shah,S., Takei,Y., Zhou,W., Lubeck,E., Yun,J., Eng,C.L., Koulena,N., Cronin,C., Karp,C., Liaw,E.J. *et al.* (2018) Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell*, **174**, 363–376.
39. Bothma,J.P., Norstad,M.R., Alamos,S. and Garcia,H.G. (2018) LlamaTags: a versatile tool to image transcription factor dynamics in live embryos. *Cell*, **173**, 1810–1822.
40. Wu,B., Miskolci,V., Sato,H., Tutucci,E., Kenworthy,C.A., Donnelly,S.K., Yoon,Y.J., Cox,D., Singer,R.H. and Hodgson,L. (2015) Synonymous modification results in high-fidelity gene expression of repetitive protein and nucleotide sequences. *Genes Dev.*, **29**, 876–886.
41. Ochiai,H., Sugawara,T. and Yamamoto,T. (2015) Simultaneous live imaging of the transcription and nuclear position of specific genes. *Nucleic Acids Res.*, **43**, e127.
42. Suter,D.M., Molina,N., Gatfield,D., Schneider,K., Schibler,U. and Naef,F. (2011) Mammalian genes are transcribed with widely different bursting kinetics. *Science*, **332**, 472–474.
43. Bahar Halpern,K., Tanami,S., Landen,S., Chapal,M., Szlak,L., Hutzler,A., Nizhberg,A. and Itzkovitz,S. (2015) Bursty gene expression in the intact mammalian liver. *Mol. Cell*, **58**, 147–156.
44. Lim,B. (2018) Imaging transcriptional dynamics. *Curr. Opin. Biotechnol.*, **52**, 49–55.
45. Fukaya,T., Lim,B. and Levine,M. (2016) Enhancer control of transcriptional bursting. *Cell*, **166**, 358–368.
46. Shermoen,A.W. and O'Farrell,P.H. (1991) Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. *Cell*, **67**, 303–310.
47. Liu,J. and Ma,J. (2013) Uncovering a dynamic feature of the transcriptional regulatory network for anterior-posterior patterning in the Drosophila embryo. *PLoS One*, **8**, e62641.
48. Kaufman,P.D. and Rando,O.J. (2010) Chromatin as a potential carrier of heritable information. *Curr. Opin. Cell Biol.*, **22**, 284–290.
49. Heard,E. and Martienssen,R.A. (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, **157**, 95–109.
50. Francis,N.J. and Kingston,R.E. (2001) Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell Biol.*, **2**, 409–421.
51. Nagashima,R., Hibino,K., Ashwin,S.S., Babokhov,M., Fujishiro,S., Imai,R., Nozaki,T., Tamura,S., Tani,T., Kimura,H. *et al.* (2019) Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *J. Cell Biol.*, **218**, 1511–1530.
52. Tumbar,T., Sudlow,G. and Belmont,A.S. (1999) Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J. Cell Biol.*, **145**, 1341–1354.
53. International Human Genome Sequencing, C. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931–945.
54. Crivat,G. and Taraska,J.W. (2012) Imaging proteins inside cells with fluorescent tags. *Trends Biotechnol.*, **30**, 8–16.
55. Specht,E.A., Braselmann,E. and Palmer,A.E. (2017) A critical and comparative review of fluorescent tools for live-cell imaging. *Annu. Rev. Physiol.*, **79**, 93–117.
56. Sander,J.D. and Joung,J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.*, **32**, 347–355.
57. Dominguez,A.A., Lim,W.A. and Qi,L.S. (2016) Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.*, **17**, 5–15.
58. Barrangou,R. and Doudna,J.A. (2016) Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.*, **34**, 933–941.
59. Leonetti,M.D., Sekine,S., Kamiyama,D., Weissman,J.S. and Huang,B. (2016) A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E3501–E3508.
60. Feng,S., Sekine,S., Pessino,V., Li,H., Leonetti,M.D. and Huang,B. (2017) Improved split fluorescent proteins for endogenous protein labeling. *Nat. Commun.*, **8**, 370.
61. Kamiyama,D., Sekine,S., Barsi-Rhyne,B., Hu,J., Chen,B., Gilbert,L.A., Ishikawa,H., Leonetti,M.D., Marshall,W.F.,

Weissman,J.S. *et al.* (2016) Versatile protein tagging in cells with split fluorescent protein. *Nat. Commun.*, **7**, 11046.
62. Merkle,F.T., Neuhausser,W.M., Santos,D., Valen,E., Gagnon,J.A., Maas,K., Sandoe,J., Schier,A.F. and Eggan,K. (2015) Efficient CRISPR-Cas9-mediated generation of knockin human pluripotent stem cells lacking undesired mutations at the targeted locus. *Cell Rep.*, **11**, 875–883.
63. Khanna,N., Hu,Y. and Belmont,A.S. (2014) HSP70 transgene directed motion to nuclear speckles facilitates heat shock activation. *Curr. Biol.*, **24**, 1138–1144.
64. Wang,H., Nakamura,M., Abbott,T.R., Zhao,D., Luo,K., Yu,C., Nguyen,C.M., Lo,A., Daley,T.P., La Russa,M. *et al.* (2019) CRISPR-mediated live imaging of genome editing and transcription. *Science*, **365**, 1301–1305.
65. Lin,S., Staahl,B.T., Alla,R.K. and Doudna,J.A. (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife*, **3**, e04766.
66. Yao,X., Zhang,M., Wang,X., Ying,W., Hu,X., Dai,P., Meng,F., Shi,L., Sun,Y., Yao,N. *et al.* (2018) Tild-CRISPR allows for efficient and precise gene knockin in mouse and human cells. *Dev. Cell*, **45**, 526–536.
67. Lee,C., Lee,S., Shin,S.G. and Hwang,S. (2008) Real-time PCR determination of rRNA gene copy number: absolute and relative quantification assays with Escherichia coli. *Appl. Microbiol. Biotechnol.*, **78**, 371–376.
68. Lalande,M. (1996) Parental imprinting and human disease. *Annu. Rev. Genet.*, **30**, 173–195.
69. Oestergaard,V.H. and Lisby,M. (2017) Transcription-replication conflicts at chromosomal fragile sites-consequences in M phase and beyond. *Chromosoma*, **126**, 213–222.