Published online 5 January 2016

Nucleic Acids Research, 2016, Vol. 44, No. 8 ***e75***
doi: 10.1093/nar/gkv1533

Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci

Baohui Chen[1], Jeffrey Hu[1], Ricardo Almeida[2], Harrison Liu[3], Sanjeev Balakrishnan[4], Christian Covill-Cooke[5], Wendell A. Lim[2,6] and Bo Huang[1,7,*]

[1]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA, [2]Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, CA 94158, USA, [3]Graduate Program of Bioengineering, University of California, San Francisco, San Francisco, CA 94143, USA, [4]Department of Cell and Tissue Biology, University of California, San Francisco, San Francisco, CA 94143, USA, [5]MRC Laboratory of Molecular Cell Biology, University College London, London, WC1E 6BT, UK, [6]Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 4143, USA and [7]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA

Received August 20, 2015; Revised November 24, 2015; Accepted December 22, 2015

ABSTRACT

In order to elucidate the functional organization of the genome, it is vital to directly visualize the interactions between genomic elements in living cells. For this purpose, we engineered the Cas9 protein from *Staphylococcus aureus* (SaCas9) for the imaging of endogenous genomic loci, which showed a similar robustness and efficiency as previously reported for *Streptococcus pyogenes* Cas9 (SpCas9). Imaging readouts allowed us to characterize the DNA-binding activity of SaCas9 and to optimize its sgRNA scaffold. Combining SaCas9 and SpCas9, we demonstrated two-color CRISPR imaging with the capability to resolve genomic loci spaced by <300 kb. Combinatorial color-mixing further enabled us to code multiple genomic elements in the same cell. Our results highlight the potential of combining SpCas9 and SaCas9 for multiplexed CRISPR-Cas9 applications, such as imaging and genome engineering.

INTRODUCTION

The type II CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) systems have been adapted to enable efficient genome editing in a wide range of cultured cells and organisms [(1–8); for reviews, see (9–12)]. Its most widely used form consists of the Cas9 nuclease and a single guide RNA (sgRNA) (1–3,13) that mimics the natural hybrid of the CRISPR RNA (crRNA) and the trans-activating CRISPR RNA (tracrRNA) (14). The 3′ end of sgRNA is the crRNA:tracrRNA scaffold that interacts with Cas9 (15–18). Target recognition by the Cas9-sgRNA complex requires Watson–Crick base pairing with the 5′ end of the sgRNA (≈20 nt) as well as a short protospacer-adjacent motif (PAM) located immediately downstream of the target DNA sequence (13,19–21). The PAM sequences are diverse among orthologous CRISPR-Cas systems. In particular, the widely used *Streptococcus pyogenes* Cas9 (SpCas9) recognizes a short 5′-NGG-3′ PAM (13,22). Cas9 contains two conserved endonuclease domains, HNH and RuvC, that cleave the two strands of the target DNA, respectively (13,21). Inactivating both catalytic active sites via point mutations results in 'nuclease-dead' Cas9 (dCas9), which retains full DNA binding activity but does not cleave DNA (13,23). The reprogrammable binding capability of dCas9 has enabled more applications such as gene expression regulation (23–29), chromatin imaging (30–33) and chromatin (34) and RNA (35) pull-down.

By expressing dCas9 from *Streptococcus pyogenes* (dSpCas9) fused to a fluorescent protein and the corresponding sgRNAs, both repetitive and non-repetitive DNA sequences can be labeled and imaged. This CRISPR imaging technique has allowed live cell examination of telomere length, gene (chromosome) copy number and the dynamics of genes in interphase as well as mitosis (30). To further expand its application in investigating the functional genome organization (36), multi-color imaging capability would be instrumental. The reported collection of Cas9 orthologs, with distinct DNA binding specificity and PAM recognition (1,37–40), constitutes a large source of CRISPR-Cas9 systems for expanding targeting flexibility and simultaneous imaging of multiple genomic loci in one cell. Several orthologous CRISPR-Cas9 systems from different bac-

*To whom correspondence should be addressed. Tel: +1 415 4761866; Fax: +1 415 5141028; Email: bo.huang@ucsf.edu

terial species, including *Neisseria meningitidis* (NmCas9) and *Streptococcus thermophilus* 1 (St1Cas9), have been applied for genome editing in human cells (1,37,38). A multicolor CRISPR system has recently been reported, using three dCas9 orthologs, dSpCas9, dNmCas9 and dSt1Cas9, fused to different fluorescent proteins (32). However, both NmCas9 and St1Cas9 require longer PAMs, such as 5′-NNNNGATT-3′ for NmCas9 (38,39), which can potentially improve the targeting specificity but limit the range of sequences that Cas9 orthologs can target. Thus, it is highly desirable to explore more Cas9 orthologs to use for more robust and versatile CRISPR-based genome visualization.

Recently, a smaller Cas9 ortholog from *Staphylococcus aureus* (SaCas9), recognizing 5′-NNGRRT-3′ PAM (R represents A or G), has been effectively used for *in vivo* genome editing using single guide RNAs (41). Given the smaller size of SaCas9, it can be more easily delivered with viral expression vectors, which is critical for both basic research and therapeutic applications. Here, we show that SaCas9 can be engineered as a tool for CRISPR imaging that is as efficient and robust as SpCas9. We further perform CRISPR imaging to gain insights into the targeting specificity of SaCas9 and the determinants that influence SaCas9 DNA-binding activity, which has not yet been fully characterized. Paired with dSpCas9, we demonstrate the capability of two-color CRISPR imaging to resolve two genomic elements spaced by <300 kb, and to color-code more than two loci for simultaneous tracking. Together these results suggest that SpCas9 and SaCas9 can be co-introduced to enable efficient multiplexed CRISPR-Cas9 functionalities beyond CRISPR imaging, such as simultaneous upregulation and downregulation of gene expression.

MATERIALS AND METHODS

dCas9 constructs

The construction of dSpCas9-EGFP has been described in our previous study (30). Similar strategies have been used to build dSaCas9-EGFP and dSaCas9-mCherry constructs. Plasmid containing dSaCas9 DNA fragment was kindly provided by Dr. Feng Zhang (MIT). The DNA sequence encoding dSaCas9 protein with D10A and N580A mutations was fused with EGFP or mCherry and two copies of SV40 NLS. Using In-Fusion HD (Clontech), we cloned these fusion proteins into a lentiviral vector containing an inducible promoter P_{TRE3G} (Tet-on 3G inducible expression system, Clontech). The DNA sequence encoding dSaCas9-EGFP protein are shown in Supplementary Table S1.

sgRNA design

dSpCas9 target sequences were designed as previously described (42). To select the target sequences of dSaCas9, we searched for 5′-$GN_{(16–23)}$-NNGRRT-3′ to target the template DNA strand. NNGRRT is the PAM sequence recognized by *S. aureus* Cas9 protein which was reported by a recent study (41). To target the non-template DNA strand, we searched for 5′-AYYCNN- $N_{(16–23)}$ C-3′. The reverse-complementary sequence of $N_{(16–23)}$ C was used as guide in the sgRNA. The guide sequences of sgRNAs used in this paper are listed in Supplementary Table S2.

sgRNA cloning

For single color CRISPR imaging, sgRNAs were cloned into a lentiviral mouse U6 expression vector derived from pSico, which coexpress mCherry and a puromycin resistance cassette from a CMV promoter (43). To clone Sp_sgRNAs, forward primer (5′-ggagaaCCACCTTGTT**G**$\mathbf{N_x}$GTTTAAGAGCTATGCT GGAAACA-3′, where GN_X is the guide sequence; BstXI restriction site is underlined) and a common reverse primer (5′- ctagtaCTCGAGAAAAAAAAGCACCGACTCGG TGCCAC-3′, where the XhoI restriction site is underlined) were used to PCR the sgRNA fragments and cloned into the sgRNA vector by Quick T4 DNA ligase. Sa_sgRNAs were cloned in the same way using the following primers: Forward primers_ 5′- ggagaaCCACCTTGTT**G**$\mathbf{N_x}$GT TTTAGTACTCTGGAAACAGAATC-3′ and a common reverse primer_5′- ctagtaCTCGAGAAAAAAAA TCTCGCCAACAAGTTG-3′). Forward primers (5′-ggagaaCCACCTTGTT**G**$\mathbf{N_x}$GTTATAGTACTCTG GAAACAGAATC -3′) were used to clone optimized Sa_sgRNAs to increase sgRNA expression. For two-color CRISPR imaging, mTagBFP was used to replace the mCherry tag in the sgRNA vector. A second sgRNA fragment (e.g. Sa_sgRNA) together with the human U6 promoter sequence was clone into the lentiviral vector which carried a sgRNA (e.g. Sp_sgRNA) expressed from a mouse U6 promoter. Thus, two sgRNAs (Sp and Sa) were expressed driven by two different U6 promoters in a same vector. The sequence of the dual-sgRNA expression system is shown in Supplementary Table S3. According to our results, mouse U6 promoter can drive sgRNA expression at a slightly higher level than human U6 promoter.

Cell culture

Human embryonic kidney (HEK) cell line HEK293T was maintained in Dulbecco's modified Eagle medium (DMEM) with high glucose (UCSF Cell Culture Facility) in 10% Tet System Approved FBS (Clontech). Human retinal pigment epithelium (RPE) cells were maintained in DMEM with GlutaMAX1 (Life Technologies) in 10% Tet System Approved FBS. Two types of RPE cells were used in this study. One was purchased directly from American Type Culture Collection (ATCC, CRL-4000), which ATCC reported as a near-diploid karyotype, with a 2N chromosome number of 46. Another aneuploid RPE cell line, which has been confirmed in our previous study (30), contains up to 4 copies of each chromosome, with a total of 73 chromosomes at 2N. All cells were grown at 37°C and 5% CO_2 in a humidified incubator.

Lentiviral production

HEK293T cells were plated into 6-well plates one day prior to transfection. 110 ng of pMD2.G plasmid, 890 ng of pCMV-dR8.91 and 1000 ng of the lentiviral vector (Tet-on 3G, dCas9-EGFP/mCherry or sgRNA) were cotransfected into HEK293T in each well using FuGENE (Promega) following the manufacturer's recommended protocol. Virus was harvested 48 h after transfection.

Generation of dual-dCas9 stable cell line

For two-color CRISPR imaging, we first infected RPE cells with dSpCas9-EGFP and Tet-on 3G lentiviruses. We then isolated clonal cell lines which expressed dSpCas9-EGFP at a suitable level to perform CRISPR imaging with high labeling efficiency and low background signal. A more detailed protocol to isolate clonal cell lines of dCas9-FP has been described (42). Next, we infected a selected dSpCas9-EGFP clonal cell line with the dSaCas9-mCherry lentivirus. Variation in dSaCas9-mCherry basal expression level was relatively small, which may be attributed to clonal Tet-on 3G transactivator expression. Thus, a second round of single cell clone selection is optional. An alternative strategy is to infect the cells with dSpCas9-EGFP, dSaCas9-mCherry and Tet-on 3G lentiviruses simultaneously and then isolate a clonal cell line with suitable expression levels of the two Cas9 proteins.

sgRNA expression

One day before transduction, dCas9-FP expressing cells were seeded into 8-well chambered coverglass (Lab-Tek II). The cells were infected by sgRNA lentivirus (by 1:6, 1:3 or 1:2 dilution) supplemented with 5 μg/ml polybrene (#TR-1003-G, Millipore) in each well. The virus dilution factor is dependent on the sgRNA efficacy and the copy number of target sites. In general, high sgRNA lentivirus dosage achieves better labeling efficiency. After 12 h, the virus-containing medium was replaced by fresh growth medium without phenol-red. 48 h post-transduction, the cells are ready for imaging. All imaging experiments were performed without doxycycline induction.

Northern blotting

Total RNA was extracted from RPE cells transduced with different sgRNA expression constructs using TRIzol (Thermo Fisher Scientific) and Direct-zol RNA Miniprep kit (Zymo Research) following manufacturer's instructions. 5 μg of total RNA samples were resolved on Novex 10% TBE-Urea PAGE gels (Life Technologies) in 0.5X TBE buffer at 120V. Equal sample loading was confirmed by staining gels with SYBR Safe prior to electroblotting (5S rRNA, 120 nt). Samples on gel were transferred to Hybond NX membranes (GE Healthcare) in 0.5X TBE for 1.5 h at 250 mA using a Mini Protean Tetra Cell apparatus (Bio-Rad) and UV crosslinked on a Stratalinker (Stratagene) twice at 120 $\mu J/cm^2$. Membranes were probed with a 5′-^{32}P-labeled DNA oligonucleotide 5′- GATAAACACGGCATTTTGCCTT-3′ diluted in modified Church-Gilbert buffer (0.5 M phosphate pH 7.2, 7% (w/v) SDS, 10 mM EDTA) with overnight incubation at 42°C. Blots were washed 2X for 30 min at 50°C in 2X SSC, 0.2% SDS before exposure with a storage phosphoscreen (GE Healthcare). A negative control RNA sample lacking the sgRNA expression cassette gave no detectable probe hybridization. Images were obtained on a Typhoon 9410 scanner (GE Healthcare) after exposure durations of 4 h to overnight. sgRNA expression levels were calculated from band intensities measured with ImageJ (44).

Identification of chromosome-specific repeats

The Tandem Repeat Finder bioinformatics tools (45) was used to identify tandem repeats in the genome. The human reference genome hg19 was downloaded from the UCSC genome browser (genome.ucsc.edu) for analysis. Unique, highly conserved repeats (Percent matches > 70) with copy number >30 located in a defined site but <5 copies in all other places of the genome were selected as candidates. We further selected candidates with >10 targetable sequences and designed multiple sgRNAs in order to ensure high CRISPR imaging efficiency. The following regions were selected for CRISPR imaging in this study: Ch1R (chromosome 1:1011233–1014805); Ch2R (chromosome 2: 242744740–242749183); Ch3R (chromosome 3: 195502236–195504816); Ch7R (chromosome 7: 158122661–158135328); Ch15R (chromosome 15: 101094498–101098864); Ch17R (chromosome 17: 1642092–1645402); Ch19R (chromosome 19: 59050388–59054262); Ch22R (chromosome 22: 49661285–49662360). Details of all guide sequences for dSpCa9 or dSaCas9 targeting are listed in Supplementary Table S2.

Fluorescence microscopy

Single-color CRISPR data were acquired on a Nikon Ti-E inverted wide-field fluorescence microscope equipped with an LED light source (Excelitas X-Cite XLED1), a 100x NA 1.40 PlanApo oil immersion objective, an motorized stage (ASI) with stage incubator (Tokai Hit), and an sCMOS camera (Hamamatsu Flash 4.0). Two-color CRISPR imaging were performed at UCSF Nikon Imaging Center using a Nikon Ti-E inverted wide-field fluorescence microscope equipped with an LED light source (Lumincor Spectra X), a 100x NA 1.40 PlanApo oil immersion objective, a motorized stage (ASI) and an sCMOS camera (Andor Zyla 5.2). Acquisitions were controlled by MicroManager. All images were taken as Z stacks at 0.4 μm steps and with a total of 15 steps. Maximum intensity projects were displayed in the figures and supplementary figures. Imaging data were analyzed with ImageJ (44) or CellProfiler (46).

RESULTS

CRISPR imaging using SaCas9

To engineer SaCas9 for visualizing specific genomic sequences, we tagged the nuclease-deficient SaCas9 (dSaCas9) (containing D10A and N580A mutations) with EGFP and two copies of nuclear localization signal (NLS) sequences. To reduce the background and cell heterogeneity, we generated clonal human retina pigment epithelial (RPE) cell lines expressing dSaCas9-EGFP at the basal level of the inducible Tet-on 3G system (Supplementary Figure S1A). Similar to dSpCas9, dSaCas9-EGFP without sgRNA was enriched in the nucleolus (Supplementary Figure S1B), suggesting that apo dCas9 (not bound to sgRNA) may nonspecifically interact with ribosomal RNA or nucleolar proteins. In contrast, lentiviral expression of sgRNAs targeting either telomeres or a tandem repeat in exon 2 of *MUC4* gene, *MUC4E* (30,47) (Figure 1A), led to fluorescence puncta at telomeres or *MUC4* loci, with a

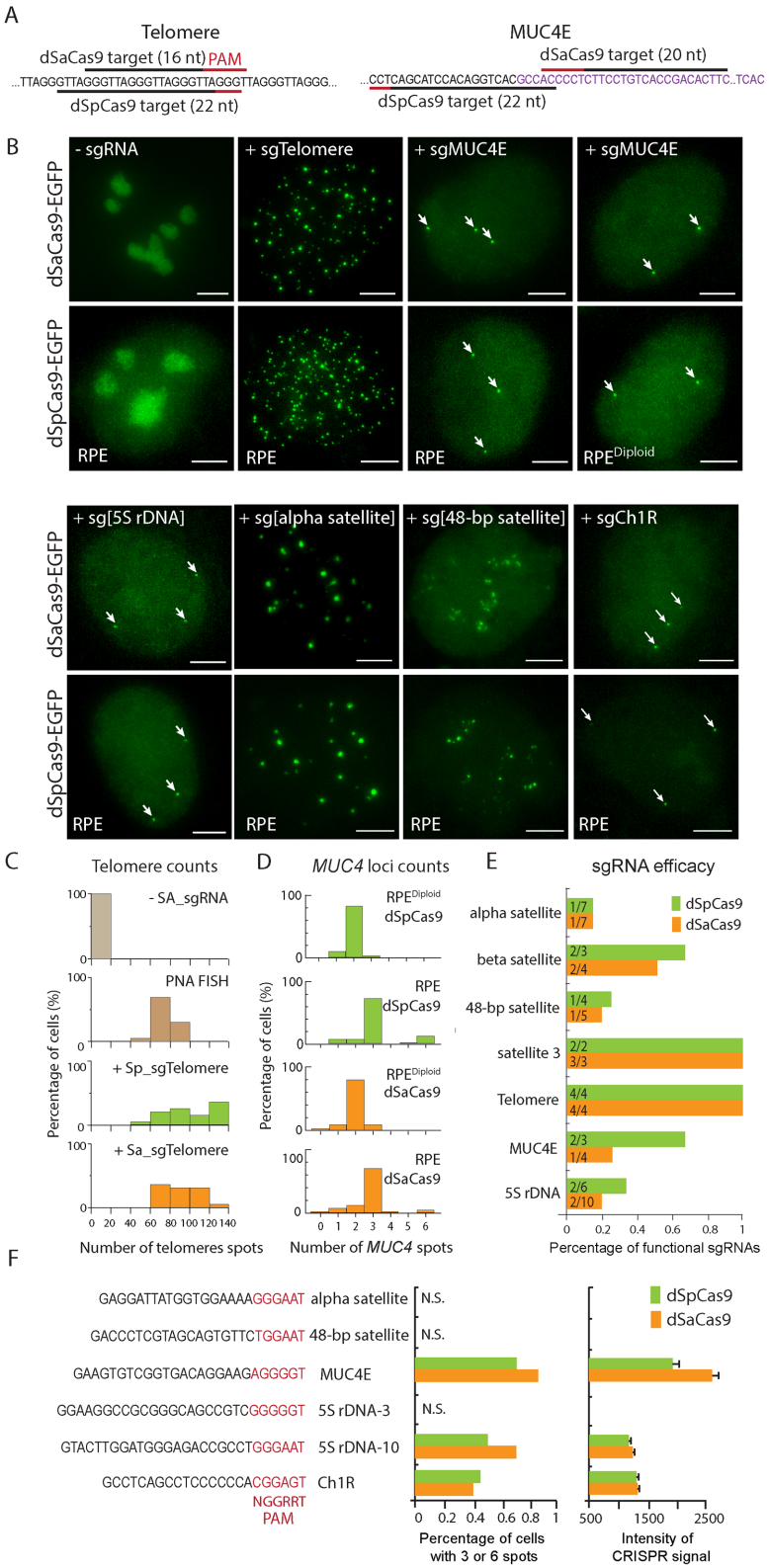

Figure 1. Comparison of SaCas9 and SpCas9 systems for CRISPR imaging. (**A**) sgRNA designs for targeting human telomeres and MUC4E. (**B**) CRISPR labeling of telomere, *MUC4* gene, 5S rDNA, centromere alpha satellite, centromere 48 bp satellite and a chromosome 1 tandem repeat in aneuploid RPE cells, and *MUC4* in diploid RPE ($RPE^{Diploid}$) cells. All images are maximum intensity projections from z stacks. Scale bars: 5 μm. (**C**) Histograms of telomere counts per cell detected by PNA FISH or the two Cas9s ($n = 20$). (**D**) Histograms of *MUC4* loci counts by CRISPR labeling in the two RPE cells. At least 25 cells were analyzed for each case. (**E**) Comparing the robustness of dSpCas9 and dSaCas9 systems as the fraction of functional sgRNAs for each locus. (**F**) Direct comparison of dSpCas9/dSaCas9 labeling efficiency by targeting the same DNA sequences. The PAM that allows targeting by both Cas9 systems are indicated in red. Labeling efficiency is defined by quantifying the percentage of cells ($n = 30$ cells) containing 3 or 6 CRISPR spots and by measuring the signal intensity of CRISPR spots ($n = 55$ spots). Error bars report ±SEM.

corresponding loss of dSaCas9-EGP signal at the nucleolus (Figure 1B). This reduction of nucleolar dCas9 signal is likely due to the assembly of Cas9-sgRNA complex which diminishes Cas9's affinity toward nucleolar components. Thus, the distribution of dCas9-EGFP can be used as an indicator of functional Cas9-sgRNA complex assembly. Both dSaCas9 and dSpCas9 could detect a similar number of telomeres over the fluorescence background, indicating that they have a comparable efficiency (Figure 1C). Consistently, both Cas9 systems detected 3 or 6 copies of *MUC4* gene in >80% of the cells, consistent with their karyotype of chromosome 3 trisomy (30) (Figure 1D). When we switched to a diploid RPE-1 cell line (32), 2 copies of *MUC4* gene were identified in both cases (Figure 1B).

Next, to assess the number of target sites that are minimally required for the efficient detection of a given genomic locus, we examined the imaging of a number of tandem repetitive loci with 10 to 90 copies of perfect repeats using either dSaCas9 or dSpCas9 (Supplementary Figure S2). We were able to detect as few as 17 copies of exact repeats by dSaCas9 in the tandem array of 5S ribosomal DNA (5S rDNA) on chromosome 1 (48), and 18 copies of exact repeats by dSpCas9 in another locus on chromosome 22 (Ch22R). Notably, we found no correlation between the copy number of exact repeats and the signal-to-background ratio, defined as the ratio between the peak intensity of fluorescent puncta at the target sites to the intensity of diffusive dCas9-EGFP signal from the nucleus. For example, a chromosome 3 locus (Ch3R) containing 70 copies of exact repeats generates lower signal compared to several other loci with fewer exact repeats. These findings illustrate that CRISPR imaging efficiency is not simply determined by the copy number of target sites. Indeed, many studies have revealed that molecular features of target sequences can influence sgRNA stability, activity and loading into Cas9 (49–53). In addition, chromatin accessibility is suggested to be a major determinant of dCas9 binding *in vivo* (54–56).

Comparison of SaCas9 and SpCas9 on DNA-labeling

In CRISPR-Cas9 gene editing, the editing efficiency varies with the target sequence. As the result, multiple sgRNAs are often designed and tested for each target loci to ensure successful editing. We can then use the fraction of successful (functional) designs as a robustness metric. In imaging applications, the number of loci labeled and the signal brightness further represent the Cas9 DNA-binding activity. Using these metrics, we compared SaCas9 and SpCas9 in imaging a variety of targets: telomeres, *MUC4E*, 5S rDNA, centromeric repetitive DNAs (including alpha, beta and 48-bp satellite DNAs) and a unique tandem repeat on chromosome 1 (Ch1R) (Figure 1B). We found that the functional fraction of target sequences was very close between dSaCas9 and dSpCas9 (Figure 1E). Because the composition of target sequence can influence sgRNA efficacy, we chose 6 sites with PAM sequence recognizable by both SaCas9 and SpCas9 ('NGGRRT') to precisely compare their targeting activity on exactly the same sequence. We found that three of these sites can be labeled by both Cas9s, whereas the other three can be labeled by neither one. For the three functional sites, the fraction of cells showing full labeling (3 or 6 spots for *MUC4*, 5S rDNA and Ch1R) and the average intensity of fluorescent puncta were similar between the two Cas9s (Figure 1F). Taken together, these data demonstrated that dSaCas9 and dSpCas9 can label genomic DNA with very similar robustness and efficiencies.

Characterization of the target sequence requirement for SaCas9

Previous studies of SaCas9-mediated genome editing have reported the optimal guide length and PAM sequence for efficient DNA cleavage (41). However, the determinants that influence dSaCas9 DNA-binding activity and specificity remains largely unknown. Thus, we characterized the target sequence requirement for dSaCas9 imaging. We first examined the guide length requirement by targeting *MUC4E* and a repeat on chromosome 19 (Ch19R) with guide lengths from 5 to 22 nucleotides (Figure 2A, Supplementary Figure S3A). Using the same metrics as described earlier, we found that dSaCas9 achieved similar labeling efficiency with 11–21 nt guides for *MUC4* and 11–22 nt for Ch19R. Although the 9 nt guide generated the same number of *MUC4* spots, the fluorescence intensity per spot was reduced by about 50%, which may be due to the unstable binding of Cas9-sgRNA to the target DNA. Despite the short guide length, we did not observe off-target labeling, which might be attributed to the requirement of multiple copies of dCas9 at the target loci to create a detectable signal. When the guide length targeting *MUC4* is further shortened to 5 nt, no specific spots representing *MUC4* loci were detected. Instead, dSaCas9 was mostly enriched in the nucleolus, potentially caused by poor Cas9-sgRNA complex formation with this truncated sgRNA. Our results thus suggest a minimum guide length of 11 nt for efficient SaCas9 DNA-binding activity. Next, we characterized the PAM recognition efficiency of SaCas9. We observed that sites with NNGRRT are more efficiently labeled than those with NNGRRG (Figure 2B), which is consistent with previous result from gene editing (41). Thus, NNGRRT should be the preferred PAM for SaCas9 sgRNA (Sa_sgRNA) design.

A series of studies showed that the targeting specificity of SpCas9 is mainly determined by the 'seed sequence' (57,58), a stretch of ≈8–12 nt at the 3′ end of the target sequence (PAM-proximal) (1,5,13,22). To assess the Cas9 specificity, several groups have created variant sgRNAs bearing one to four nucleotide mismatches in the complementary region and then examined the abilities of these sgRNAs to direct Cas9 activity (1,13,59–61). We used the same approach to explore SaCas9's DNA-binding specificity. We found that a mismatch at the first base from the PAM completely abolished CRISPR spots targeting both *MUC4* and Ch19R. For *MUC4* loci targeting, single nucleotide mismatches at positions 3, 5, 7, 9 and 11 could be tolerated but substantially reduced the imaging efficiency, whereas even multiple mismatches further away from the PAM did not affect the labeling efficiency (Figure 2C). For Ch19R targeting, similar effects were observed (Supplementary Figure S3B). We further examined *MUC4* labeling efficiency using 27 mismatched sgRNAs, containing all possible single-nucleotide in positions 1, 3, 5, 7, 9, 11, 13, 15 and 19 (Figure 2D). We observed that the degree of tolerance varied with the iden-

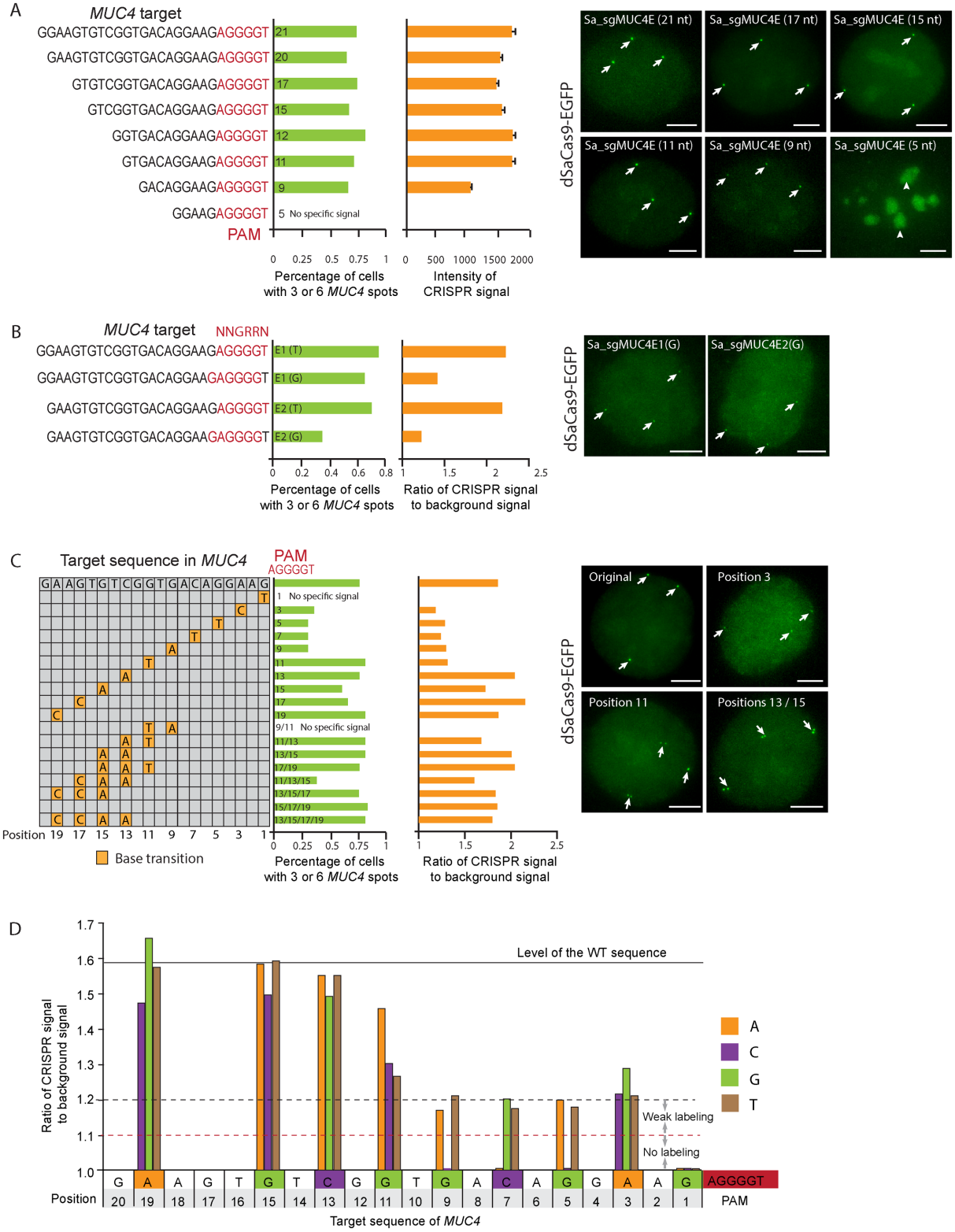

Figure 2. Characterization of dSaCas9 DNA-binding activity. (**A**) Effect of sgRNA guide length on the labeling efficiency of MUC4E. The labeling efficiency was measured by the percentage of cells ($n \geq 30$ cells) containing 3 or 6 CRISPR spots and the signal intensity of CRISPR spots ($n \geq 55$ spots). Error bars report ±SEM. (**B**) Comparison of dSaCas9 labeling efficiency with NNGRRT and NNGRRG PAMs. The labeling efficiency was measured by the percentage of cells ($n \geq 30$ cells) containing 3 or 6 CRISPR spots and the ratio of CRISPR signal to the background signal ($n = 60$ spots). (**C**) Effect of sgRNA–DNA mismatches on dSaCas9 activity. The positions and identities of base transition are indicated in yellow. The labeling efficiency was determined in the same way as in (B). **(D)***MUC4* labeling efficiency illuminated by signal-to-background ratio after transformation of all possible nucleotides at each position within *MUC4* target sequence. 20 nucleotides of the protospacer sequence is shown. A least 60 CRISPR spots are measured. All images are maximum intensity projections from z stacks. Scale bars: 5 μm.

tity of a particular mismatch in the PAM-proximal region (positions 1, 3, 5, 7, 9 and 11) but is consistently high at the PAM-distal region (positions 13, 15 and 19). These results indicate that SaCas9 DNA-binding activity is sensitive to sgRNA-DNA pairing in the first 11 nucleotides, but may partially tolerate single-base mismatches dependent on the identity of mismatched bases. The seed sequence length we identified by direct visualization of dSaCas9 targeting is longer than that by chromatin immunoprecipitation (ChIP) (7–8 nt) (41). A broader sample of targets and different sgRNAs need to be analyzed for a more comprehensive understanding of SaCas9 targeting specificity.

Improvement of the sgRNA scaffold for SaCas9

The sgRNA was constructed by fusing the crRNA and tracrRNA, with different scaffold designs resulting in different targeting efficiencies (10,30,60). For example, the sgRNA for SaCas9 (Sa_sgRNA) was optimized for DNA cleavage activity by shortening the length of crRNA:tracrRNA repeat/anti-repeat stem loop included in the resulting fusion RNA (41). However, we frequently observed strong accumulation of dSaCas9-EGFP in the nucleolus using this scaffold in our imaging system. Similar to dSpCas9, increasing the dosage of sgRNA-expressing lentivirus reduced the nucleolar signal from dSaCas9-EGFP and increased the labeling efficiency (Supplementary Figure S4). Considering the low expression level of dSaCas9-GFP, this result suggests that sgRNA is the limiting factor in imaging efficiency. To further improve the sgRNA scaffold, we removed a putative RNA Pol III terminator sequence (4 consecutive U's) (62) from the stem loop via A-U flip (Figure 3A). We performed Northern blot to measure sgRNA abundance in cells. sgRNAs level of all four different A-U flip configurations were found to be ≈5-fold higher than the original Sa_sgRNA scaffold (Figure 3B), suggesting that poly-U tract in the stem loop might induce premature termination of sgRNA transcription. We then directly compared the labeling efficiency of these A-U flip modifications to the original sgRNA scaffold design. Imaging of telomeres and 48 bp centromeric satellite showed that flipping the 3rd A-U pair increased telomere puncta number and substantially decreased the nucleolar signal most dramatically (Figure 3C, D and E). We also observed mild effects with other A-U flips, which might be due to their different effects on Cas9-sgRNA complex formation or DNA targeting. We further confirmed the improvement in labeling with the 3rd A-U flip by testing two more sgRNAs to label beta satellite DNA and 5S rDNA, which can only be weakly labeled with the original design (Supplementary Figure S5). It is important to note that the nucleolar signal was also greatly reduced with the optimized sgRNA, creating a clean background for genomic loci detection. We expect this improved sgRNA design will benefit CRISPR applications with SaCas9 not only for imaging but also for gene editing and gene expression regulation.

Dual-color CRISPR imaging using dSaCas9 and dSpCas9

To perform dual-color imaging, we used a lentiviral vector to co-express Sp and Sa_sgRNAs using mouse and human U6 promoters (Supplementary Figure S6A). We generated a stable cell line expressing both dSpCas9-EGFP and dSaCas9-mCherry under Tet-on 3G promoter (Figure 4A, Supplementary Figure S6B). With two sgRNAs targeting the same loci (telomere or *MUC4E*), we observed perfect co-localization of EGFP and mCherry puncta (Figure 4B, Supplementary Figure S7A), demonstrating their compatibility for simultaneous imaging. When targeting well-isolated loci (*MUC4E* and 5S rDNA, or *MUC1* exon 1 and *MUC4E*), we observed no cross-reactivity (Figure 4B, Supplementary Figure S7B). To demonstrate the capability of dual color CRISPR imaging to study the physical interaction between genomic elements, we labeled *MUC4E* and another unique repetitive region at about ≈272 kb upstream of *MUC4,* denoted as $MUC4E^{-272kb}$, or *MUC4E* and a tandem repeat in intron 3 of *MUC4* gene, *MUC4I*. We were able to resolve the pair of two spots from *MUC4E* and $MUC4E^{-272kb}$, whereas spots from *MUC4E* and *MUC4I*, with their mid-point separated by ≈7 kb in the DNA sequences, overlap partially (Figure 4C). Finally, we confirmed that dSaCas9 and dSpCas9 have very similar efficiencies when both proteins are present (Figure 4D, E), and expressing both sgRNAs from the same vector is similar in efficiency to using two separate sgRNA vectors (Figure 4F). Together, these experiments indicate that SpCas9 and SaCas9 can be applied to efficiently target distinct DNA sequences in the same cell.

Color-coding multiple genomic elements by dSpCas9 and dSaCas9

With two orthogonal colors for CRISPR imaging, we can further expand the palette size through color mixing (Figure 5A). As an example, in the same RPE cells we labeled 5S rDNA on chromosome 1 and *MUC4I* on chromosome 3 both with dSpCas9-EGFP as well as the chromosome 17 tandem repeat *Ch17R* and *MUC4E* with dSaCas9-mCherry. The four sgRNAs were expressed from two vectors as indicated (Figure 5B). In the two-color image, we were able to identify 5S rDNA as 3 green foci, *Ch17R* as 3 red foci and *MUC4* gene as 3 pairs of overlapping green and red foci (Figure 5B). In a second case, we labeled three loci from three different chromosomes (Chromosomes 1, 7 and 19) with dSpCas9-EGFP, dSaCas9-mCherry and both, respectively (Figure 5C). The number of green, red and yellow (overlapping green and red) spots in the two-channel image exactly matched the chromosome copy number from our previous karyotype analysis of this cell line (30). Collectively, the similarity of SaCas9 and SpCas9 systems in labeling efficiency and PAM flexibility makes them an ideal pair of orthogonal Cas9/sgRNA combinations for two-color imaging, while our color mixing scheme expands the number of loci that can be monitored simultaneously.

DISCUSSION

In this study, we characterized the *S. aureus* CRISPR-Cas9 system for CRISPR imaging. Specifically, we used a two-step procedure to express dCas9-EGFP and sgRNA: first generating a clonal dCas9-EGFP stable cell line with optimal dCas9 expression level, and then expressing sgRNA

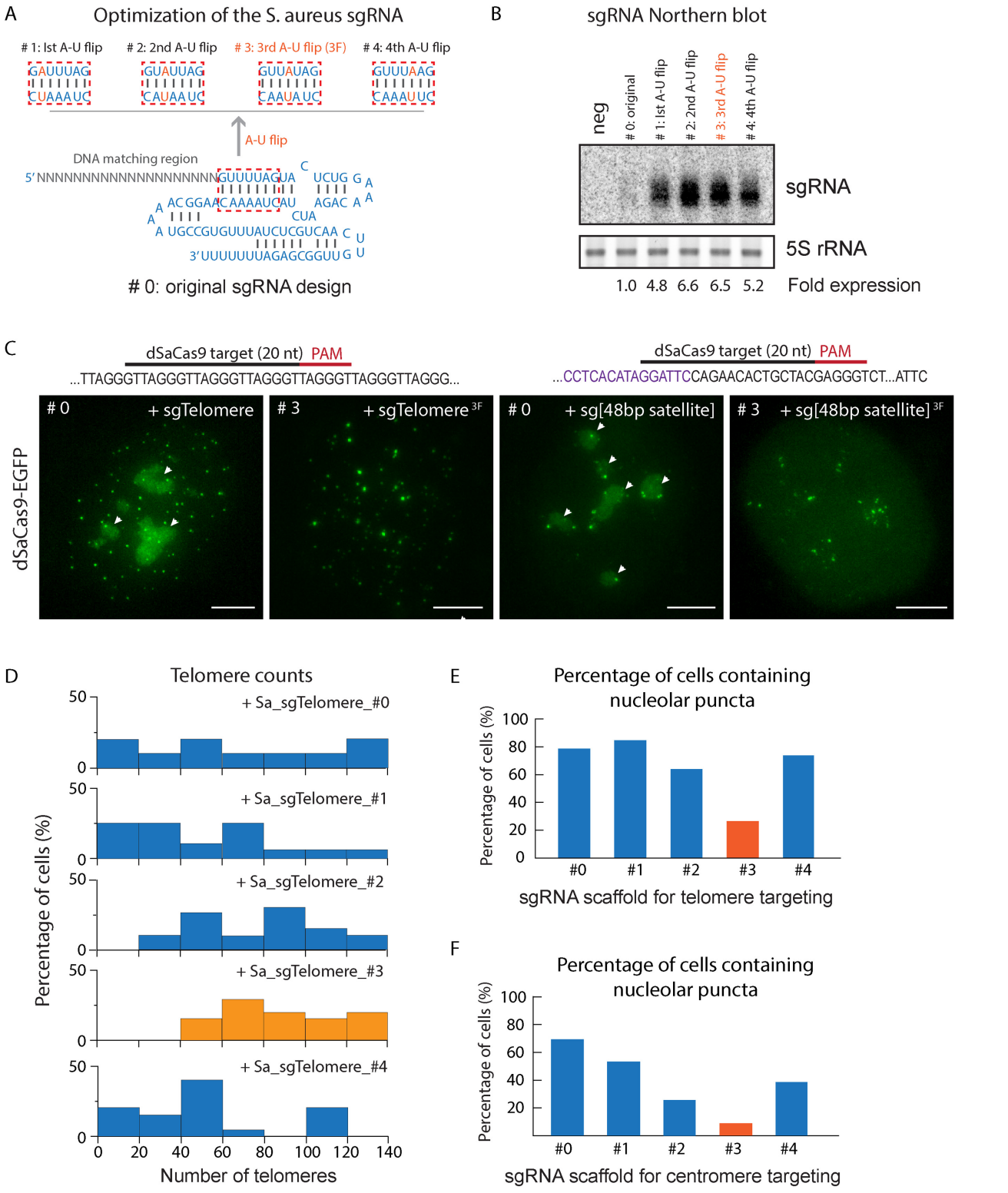

Figure 3. Optimization of sgRNA scaffold for SaCas9. (**A**) Sequences of the original and modified sgRNA scaffolds with A-U pair flip at different positions. (**B**) Northern blotting to detect *S. aureus* sgRNA abundance in RPE cells. Total RNA samples from cells expressing either wild-type or A-U flipped sgRNAs were analyzed by Northern blot using a sgRNA-specific probe to determine *in vivo* steady-state sgRNA levels. Band intensities relative to wild-type sample are shown below. 5S rRNA was imaged by gel staining as loading control. sgRNA design for targeting human telomeres and 48-bp family satellite DNA using dSaCas9. (**C**) CRISPR imaging of telomeres and 48-bp satellite DNA by dSaCas9-EGFP with original and optimized sgRNA scaffold. Arrowheads: nucleolar puncta. Scale bars: 5 μm. (**D**) Histograms of telomere counts per cell detected by dSaCas9-EGFP with the original and four modified sgRNAs ($n = 20$ cells). (**E**) Percentage of cells contain nucleolar puncta with different sgRNA scaffolds targeting telomeres. $n \geq 50$ cells. (**F**) Percentage of cells containing nucleolar puncta with different sgRNA scaffolds targeting 48-bp satellite DNA. $n \geq 58$ cells.

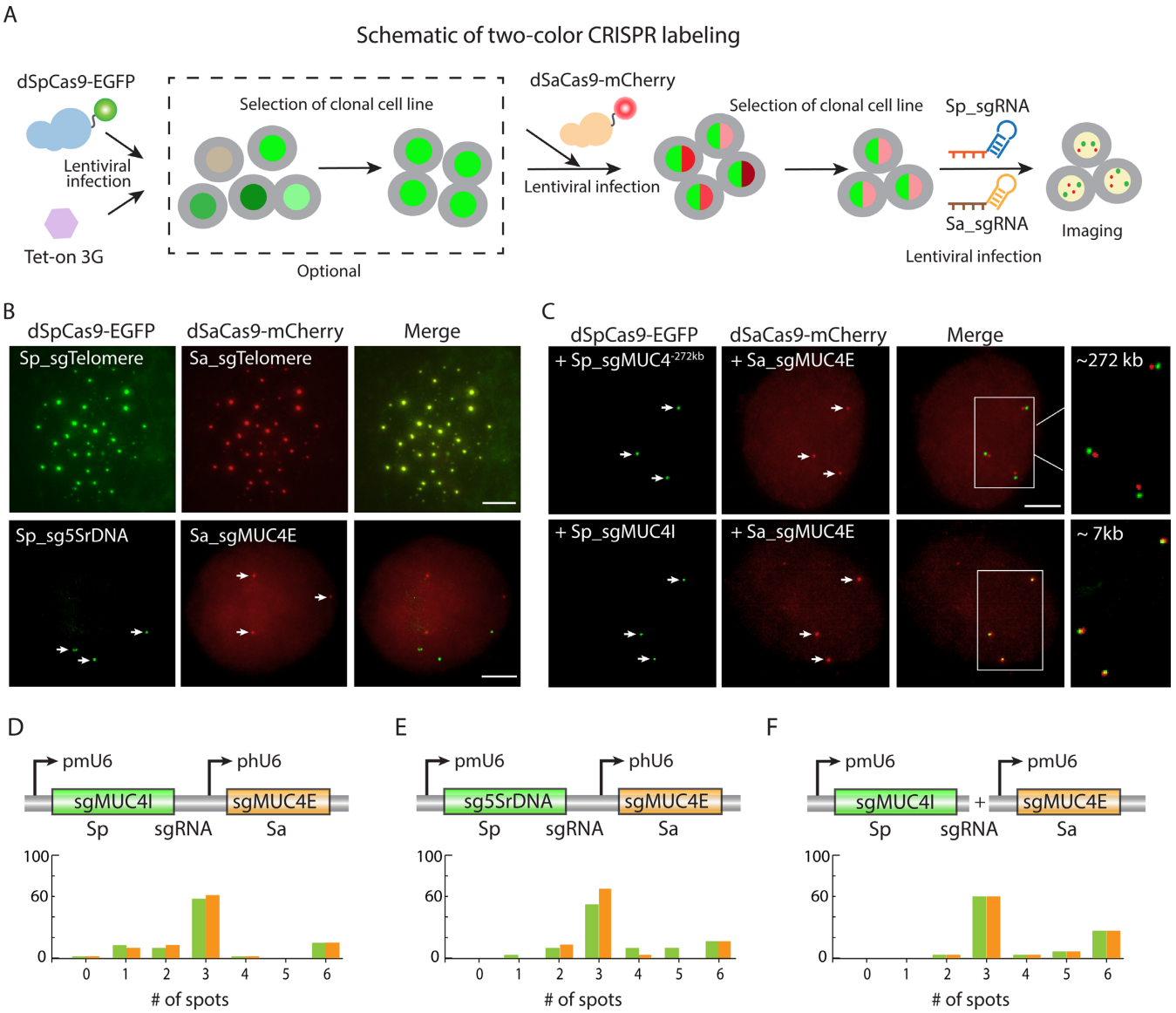

Figure 4. Two-color CRISPR imaging using dSaCas9 and dSpCas9. (**A**) Schematic of two-color CRISPR imaging. (**B**) Co-labeling of human telomeres or two loci on different chromosomes (5S rDNA and *MUC4*) by dSpCas9-EGFP and dSaCas9-mCherry in the same cell. (**C**) Simultaneous labeling of two close-by loci, separated by ≈272 kb and ≈7 kb, respectively. All images are maximum intensity projections from z stacks. Scale bars: 5 μm. (**D**) Histograms of the counts of dSpCas9-EGFP spots indicating MUC4I and dSaCas9-mCherry spots indicating MUC4E. Sp_sgMUC4I controlled by mouse U6 promoter and Sa_sgMUC4E controlled by human U6 promoter were expressed from a same vector. (**E**) Histograms of the counts of dSpCas9-EGFP spots indicating 5S rDNA loci and dSaCas9-mCherry spots indicating MUC4E. Sp_sg5S rDNA and Sa_sgMUC4E were also expressed from a same vector driven by two U6 promoters. (**F**) Histograms of the counts of dSpCas9-EGFP spots indicating MUC4 intron region (MUC4I) and dSaCas9-mCherry spots indicating *MUC4* exon region (MUC4E). Two sgRNAs were expressed from two separate vectors, both driven by mouse U6 promoter. Aneuploid RPE cells with three copies of chromosome 1 (where *MUC4* is localized) and chromosome 3 (where 5S rDNA is localized) were used to perform all the CRISPR labeling experiments in this figure.

from lentiviral vectors. This procedure provides an ideal platform with minimal Cas9 expression variation in order to characterize and compare the efficacy of sgRNAs. On the other hand, CRISPR imaging by transient co-expression of dCas9 and sgRNA is faster and easier, but the labeling efficiency can be relatively low and varies substantially across cell populations (data not shown).

Increasing sgRNA expression can drive dCas9 translocation from nucleolus to nucleoplasm, leading to higher targeting efficiency (Supplemental Figure S4). However, dSaCas9 is enriched in nucleolus even with co-expression of some guide RNAs, suggesting that efficient dSaCas9-sgRNA complex formation is counteracted by an intrinsic affinity of dSaCas9 to nucleolar proteins or RNA. The bright nucleolar puncta can interfere with adequate imaging of labeled genomic loci. Therefore, controlling sgRNA abundance through enhanced expression and stability is particularly useful for CRISPR imaging. We have improved Sa_sgRNA by removing the poly-U tract in the 5′ scaffold region, similarly to the strategy successfully applied

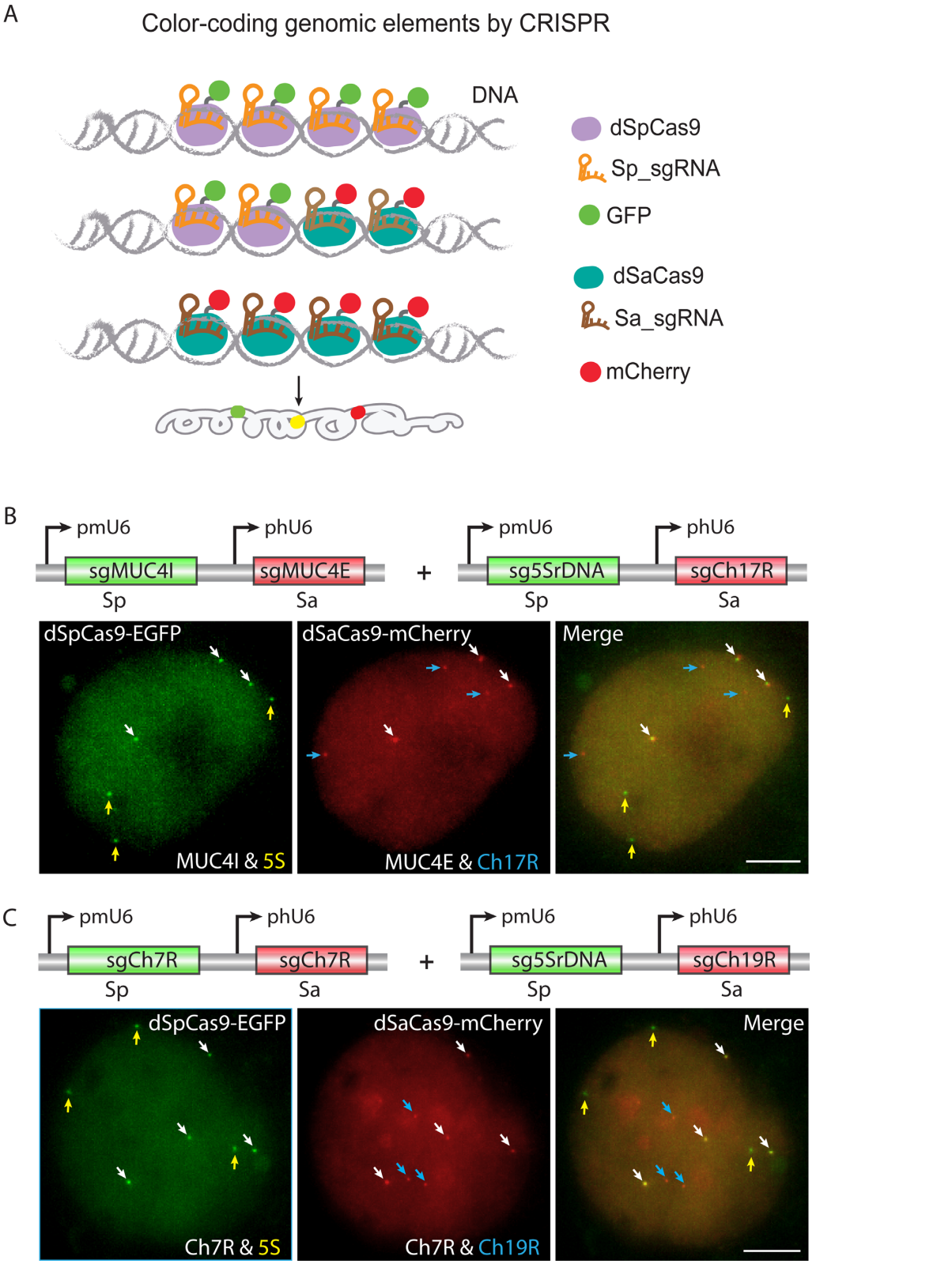

Figure 5. Color-coding of genomic elements. (**A**) Schematic of color-coding to distinguish three genomic elements by two color labeling and imaging. (**B**) Imaging of three loci (*MUC4*, 5S rDNA and Ch17R) from chromosomes 1, 3 and 17, respectively, in the aneuploid RPE cell. The chromosome copy numbers from visual karyotyping are listed on the right. (**C**) Imaging of three loci (5S rDNA, Ch7R and Ch19R) from chromosomes 1, 7 and 19, respectively, in the aneuploidy RPE cell. Chromosomes 1 and 19 have trisomy whereas chromosome 17 has 4 copies. All images are maximum intensity projections from z stacks. Scale bars: 5 μm.

to Sp_sgRNA (30). Optimized Sa_sgRNAs enable more efficient labeling, accompanied by significant reductions of noise from nucleolus. Together with other studies (54,63), we suggest that poly-U tracts should also be avoided in the guide portion of the sgRNA.

A number of studies have demonstrated that the molecular features of target sites and the sgRNA secondary architecture influence sgRNA activity (49,52,53,63). Consequently, the efficiency of Cas9 targeting for a genomic locus is strongly affected by the sgRNA used. Consistent with previous findings, only a fraction of sgRNAs that we designed for a given locus are functional (Figure 1E). Intriguingly, all the genomic regions that can be labeled by dSpCas9 can also be detected by dSaCas9, although the percentage of active sgRNAs are different. It is notable that these two orthologs behave similarly when targeting exactly the same sequence (Figure 1F), revealing that the molecular features of target sites are indeed a critical determinants of CRISPR activity. Our results also show that a functional sgRNA can always be obtained by testing a number of sgRNAs as long as other factors, such as the number of repeats influencing signal and chromatin structure permissive to sgRNA binding, are in favor of efficient labeling by CRISPR systems.

By using imaging as a readout to determine sgRNA targeting efficiency, we suggest a minimum guide length of 11 nt for efficient DNA binding of dSaCas9. Varying guide lengths between 11 and 22 nt led to similar imaging efficiency, which is different from the optimal guide length defined by cleavage induced indels (21–23 nt) (41). Two groups have mapped genome-wide binding sites of dSpCas9 and revealed that most of sites bound by Cas9 do not result in cleavage (54,55). These findings support a proposed two-state model (54) that pairing of the seed region triggers Cas9 binding but extensive paring with target DNA is required for Cas9 cleavage, which was also demonstrated by structural and biochemical studies (17). Therefore, although Cas9 targeting characteristics has been extensively explored by genome editing studies, our results provide additional guidelines for chromatin imaging and other applications that rely on dCas9 binding, such as gene regulation by CRISPRi/a (23,61) and Cas9-based purification strategies.

We identified a seed sequence of 10–12 nucleotides in the PAM-proximal region which is critical for SaCas9 binding specificity. However, the degree of effects at each position is dependent on the base identities and the molecular features of target sequence (Figure 2D, Supplementary Figure S3B). These properties are similar to that of SpCas9 (9–12). Our findings demonstrate that sequences with mismatches in the PAM-distal region can still be labeled efficiently in cells. However, it is important to note that CRISPR imaging is less affected by off-target binding than genome editing since multiple copies of dCas9 are required at the target locus to generate a detectable signal over the background. Our characterization results may also provide guidelines for target sequence choice for other SaCas9 applications, such as genome engineering.

Orthologous CRISPR systems, with their different PAM requirements, provide a powerful platform for achieving multiplex functional outputs. While this paper was under review, a Cpf1-based CRISPR system was reported to mediate robust DNA cleavage in human cells with features distinct from Cas9 (64). An alternative to fusing the fluorescent protein to dCas9 protein is to recruit fluorescent proteins directly to the sgRNA using well-defined RNA-protein interaction pairs, which can then dramatically expand the number of separate channels while using the same dCas9 protein (28,65,66). Thus, the natural diversity of CRISPR systems, the ability to engineer Cas9/Cpf1 derivatives with altered PAM specificities (67,68) and sgRNA scaffold engineering provide a wealth of opportunities for expanding the repertoire of CRISPR-based genome imaging tools. These capabilities would enable applications such as the characterization of chromatin compaction (69), tracking of chromosome translocation events (70) and the visualization of chromatin contacts and loops (71) at endogenous, unmodified loci in living cells. It is worth noting that we have demonstrated the ability to resolve genomic elements at the scale of ≈100 kb (Figure 3C) which is comparable to the current resolution of chromatin conformation capture (particularly Hi-C) experiments (71). Therefore, we envision that dynamic CRISPR live imaging will complement classic chromosomal topology mapping approaches in understanding the relationship between three-dimensional genome organization and regulation of gene expression.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Feng Zhang and Fei Ann Ran for generously providing the SaCas9 plasmids. H.L. receives the NSF Graduate Research Fellowship. C.C. is supported by the UCL-UCSF exchange program.

FUNDING

W. M. Keck Foundation; National Institutes of Health [R33EB019784 to B.C. and B.H., U19CA179512 to S.B. and B.H]. Funding for open access charge: W. M. Keck Foundation.
Conflict of interest statement. None declared.

REFERENCES

1. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
2. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
3. Jinek,M., East,A., Cheng,A., Lin,S., Ma,E. and Doudna,J. (2013) RNA-programmed genome editing in human cells. *eLife*, **2**, e00471.
4. Cho,S.W., Kim,S., Kim,J.M. and Kim,J.S. (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.*, **31**, 230–232.
5. Jiang,W., Bikard,D., Cox,D., Zhang,F. and Marraffini,L.A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.*, **31**, 233–239.
6. DiCarlo,J.E., Norville,J.E., Mali,P., Rios,X., Aach,J. and Church,G.M. (2013) Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res.*, **41**, 4336–4343.

7. Hwang,W.Y., Fu,Y., Reyon,D., Maeder,M.L., Kaini,P., Sander,J.D., Joung,J.K., Peterson,R.T. and Yeh,J.R. (2013) Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One*, **8**, e68708.
8. Wang,H., Yang,H., Shivalila,C.S., Dawlaty,M.M., Cheng,A.W., Zhang,F. and Jaenisch,R. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, **153**, 910–918.
9. Mali,P., Esvelt,K.M. and Church,G.M. (2013) Cas9 as a versatile tool for engineering biology. *Nat. Methods*, **10**, 957–963.
10. Sander,J.D. and Joung,J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.*, **32**, 347–355.
11. Hsu,P.D., Lander,E.S. and Zhang,F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.
12. Doudna,J.A. and Charpentier,E. (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, **346**, 1258096.
13. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
14. Deltcheva,E., Chylinski,K., Sharma,C.M., Gonzales,K., Chao,Y., Pirzada,Z.A., Eckert,M.R., Vogel,J. and Charpentier,E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**, 602–607.
15. Jinek,M., Jiang,F., Taylor,D.W., Sternberg,S.H., Kaya,E., Ma,E., Anders,C., Hauer,M., Zhou,K., Lin,S. *et al.* (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, **343**, 1247997.
16. Nishimasu,H., Cong,L., Yan,W.X., Ran,F.A., Zetsche,B., Li,Y., Kurabayashi,A., Ishitani,R., Zhang,F. and Nureki,O. (2015) Crystal Structure of Staphylococcus aureus Cas9. *Cell*, **162**, 1113–1126.
17. Jiang,F., Zhou,K., Ma,L., Gressel,S. and Doudna,J.A. (2015) STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science*, **348**, 1477–1481.
18. Nishimasu,H., Ran,F.A., Hsu,P.D., Konermann,S., Shehata,S.I., Dohmae,N., Ishitani,R., Zhang,F. and Nureki,O. (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, **156**, 935–949.
19. Horvath,P., Romero,D.A., Coute-Monvoisin,A.C., Richards,M., Deveau,H., Moineau,S., Boyaval,P., Fremaux,C. and Barrangou,R. (2008) Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *J. Bacteriol.*, **190**, 1401–1412.
20. Mojica,F.J., Diez-Villasenor,C., Garcia-Martinez,J. and Almendros,C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, **155**, 733–740.
21. Gasiunas,G., Barrangou,R., Horvath,P. and Siksnys,V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E2579–E2586.
22. Sternberg,S.H., Redding,S., Jinek,M., Greene,E.C. and Doudna,J.A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, **507**, 62–67.
23. Qi,L.S., Larson,M.H., Gilbert,L.A., Doudna,J.A., Weissman,J.S., Arkin,A.P. and Lim,W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.
24. Bikard,D., Jiang,W., Samai,P., Hochschild,A., Zhang,F. and Marraffini,L.A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.*, **41**, 7429–7437.
25. Gilbert,L.A., Larson,M.H., Morsut,L., Liu,Z., Brar,G.A., Torres,S.E., Stern-Ginossar,N., Brandman,O., Whitehead,E.H., Doudna,J.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154**, 442–451.
26. Maeder,M.L., Linder,S.J., Cascio,V.M., Fu,Y., Ho,Q.H. and Joung,J.K. (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods*, **10**, 977–979.
27. Perez-Pinera,P., Kocak,D.D., Vockley,C.M., Adler,A.F., Kabadi,A.M., Polstein,L.R., Thakore,P.I., Glass,K.A., Ousterout,D.G., Leong,K.W. *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods*, **10**, 973–976.
28. Mali,P., Aach,J., Stranges,P.B., Esvelt,K.M., Moosburner,M., Kosuri,S., Yang,L. and Church,G.M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.*, **31**, 833–838.
29. Konermann,S., Brigham,M.D., Trevino,A.E., Hsu,P.D., Heidenreich,M., Cong,L., Platt,R.J., Scott,D.A., Church,G.M. and Zhang,F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, **500**, 472–476.
30. 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.
31. Anton,T., Bultmann,S., Leonhardt,H. and Markaki,Y. (2014) Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus*, **5**, 163–172.
32. Ma,H., Naseri,A., Reyes-Gutierrez,P., Wolfe,S.A., Zhang,S. and Pederson,T. (2015) Multicolor CRISPR labelling of chromosomal loci in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3002–3007.
33. Deng,W., Shi,X., Tjian,R., Lionnet,T. and Singer,R.H. (2015) CASFISH: CRISPR/Cas9-mediated in situ labelling of genomic loci in fixed cells. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 11870–11875.
34. Fujita,T. and Fujii,H. (2013) Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochem. Biophys. Res. Commun.*, **439**, 132–136.
35. O'Connell,M.R., Oakes,B.L., Sternberg,S.H., East-Seletsky,A., Kaplan,M. and Doudna,J.A. (2014) Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*, **516**, 263–266.
36. Misteli,T. (2013) The cell biology of genomes: bringing the double helix to life. *Cell*, **152**, 1209–1212.
37. Esvelt,K.M., Mali,P., Braff,J.L., Moosburner,M., Yaung,S.J. and Church,G.M. (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods*, **10**, 1116–1121.
38. Hou,Z., Zhang,Y., Propson,N.E., Howden,S.E., Chu,L.F., Sontheimer,E.J. and Thomson,J.A. (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 15644–15649.
39. Fonfara,I., Le Rhun,A., Chylinski,K., Makarova,K.S., Lecrivain,A.L., Bzdrenga,J., Koonin,E.V. and Charpentier,E. (2014) Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.*, **42**, 2577–2590.
40. Chylinski,K., Makarova,K.S., Charpentier,E. and Koonin,E.V. (2014) Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.*, **42**, 6091–6105.
41. Ran,F.A., Cong,L., Yan,W.X., Scott,D.A., Gootenberg,J.S., Kriz,A.J., Zetsche,B., Shalem,O., Wu,X., Makarova,K.S. *et al.* (2015) In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, **520**, 186–191.
42. Chen,B. and Huang,B. (2014) In: Doudna,J.A. and Sontheimer,E.J. (eds.), *The use of CRISPR/cas9, ZFNs, TALENs in generating site specific genome alterations*. Academic Press, UK, Vol. **546**, pp. 337–354.
43. Larson,M.H., Gilbert,L.A., Wang,X., Lim,W.A., Weissman,J.S. and Qi,L.S. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.*, **8**, 2180–2196.
44. Schneider,C.A., Rasband,W.S. and Eliceiri,K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
45. Benson,G. (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.*, **27**, 573–580.
46. Carpenter,A.E., Jones,T.R., Lamprecht,M.R., Clarke,C., Kang,I.H., Friman,O., Guertin,D.A., Chang,J.H., Lindquist,R.A., Moffat,J. *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.*, **7**, R100.
47. Nollet,S., Moniaux,N., Maury,J., Petitprez,D., Degand,P., Laine,A., Porchet,N. and Aubert,J.P. (1998) Human mucin gene MUC4: organization of its 5′-region and polymorphism of its central tandem repeat array. *Biochem. J.*, **332**, 739–748.
48. Stults,D.M., Killen,M.W., Pierce,H.H. and Pierce,A.J. (2008) Genomic architecture and inheritance of human ribosomal RNA gene clusters. *Genome Res.*, **18**, 13–18.

49. Doench,J.G., Hartenian,E., Graham,D.B., Tothova,Z., Hegde,M., Smith,I., Sullender,M., Ebert,B.L., Xavier,R.J. and Root,D.E. (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.*, **32**, 1262–1267.
50. Wang,T., Wei,J.J., Sabatini,D.M. and Lander,E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343**, 80–84.
51. Gagnon,J.A., Valen,E., Thyme,S.B., Huang,P., Akhmetova,L., Pauli,A., Montague,T.G., Zimmerman,S., Richter,C. and Schier,A.F. (2014) Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One*, **9**, e98186.
52. Moreno-Mateos,M.A., Vejnar,C.E., Beaudoin,J.D., Fernandez,J.P., Mis,E.K., Khokha,M.K. and Giraldez,A.J. (2015) CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods*, **12**, 982–988.
53. Chari,R., Mali,P., Moosburner,M. and Church,G.M. (2015) Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat. Methods*, **12**, 823–826.
54. Wu,X., Scott,D.A., Kriz,A.J., Chiu,A.C., Hsu,P.D., Dadon,D.B., Cheng,A.W., Trevino,A.E., Konermann,S., Chen,S. *et al.* (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.*, **32**, 670–676.
55. Kuscu,C., Arslan,S., Singh,R., Thorpe,J. and Adli,M. (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol.*, **32**, 677–683.
56. Knight,S.C., Xie,L., Deng,W., Guglielmi,B., Witkowsky,L.B., Bosanac,L., Zhang,E.T., El Beheiry,M., Masson,J.B., Dahan,M. *et al.* (2015) Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, **350**, 823–826.
57. Semenova,E., Jore,M.M., Datsenko,K.A., Semenova,A., Westra,E.R., Wanner,B., van der Oost,J., Brouns,S.J. and Severinov,K. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10098–10103.
58. Wiedenheft,B., van Duijn,E., Bultema,J.B., Waghmare,S.P., Zhou,K., Barendregt,A., Westphal,W., Heck,A.J., Boekema,E.J., Dickman,M.J. *et al.* (2011) RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10092–10097.
59. Fu,Y., Foden,J.A., Khayter,C., Maeder,M.L., Reyon,D., Joung,J.K. and Sander,J.D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, **31**, 822–826.
60. Hsu,P.D., Scott,D.A., Weinstein,J.A., Ran,F.A., Konermann,S., Agarwala,V., Li,Y., Fine,E.J., Wu,X., Shalem,O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.
61. Gilbert,L.A., Horlbeck,M.A., Adamson,B., Villalta,J.E., Chen,Y., Whitehead,E.H., Guimaraes,C., Panning,B., Ploegh,H.L., Bassik,M.C. *et al.* (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, **159**, 647–661.
62. Nielsen,S., Yuzenkova,Y. and Zenkin,N. (2013) Mechanism of eukaryotic RNA polymerase III transcription termination. *Science*, **340**, 1577–1580.
63. Wang,T., Wei,J.J., Sabatini,D.M. and Lander,E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343**, 80–84.
64. Zetsche,B., Gootenberg,J.S., Abudayyeh,O.O., Slaymaker,I.M., Makarova,K.S., Essletzbichler,P., Volz,S.E., Joung,J., van der Oost,J., Regev,A. *et al.* (2015) Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*, **163**, 759–771.
65. Zalatan,J.G., Lee,M.E., Almeida,R., Gilbert,L.A., Whitehead,E.H., La Russa,M., Tsai,J.C., Weissman,J.S., Dueber,J.E., Qi,L.S. *et al.* (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*, **160**, 339–350.
66. Shechner,D.M., Hacisuleyman,E., Younger,S.T. and Rinn,J.L. (2015) Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods*, **12**, 664–670.
67. Kleinstiver,B.P., Prew,M.S., Tsai,S.Q., Topkar,V.V., Nguyen,N.T., Zheng,Z.L., Gonzales,A.P.W., Li,Z.Y., Peterson,R.T., Yeh,J.R.J. *et al.* (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, **523**, U481–U249.
68. Kleinstiver,B.P., Prew,M.S., Tsai,S.Q., Nguyen,N.T., Topkar,V.V., Zheng,Z. and Joung,J.K. (2015) Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.*, **33**, 1293–1298.
69. Hu,Y., Kireev,I., Plutz,M., Ashourian,N. and Belmont,A.S. (2009) Large-scale chromatin structure of inducible genes: transcription on a condensed, linear template. *J. Cell Biol.*, **185**, 87–100.
70. Roukos,V., Voss,T.C., Schmidt,C.K., Lee,S., Wangsa,D. and Misteli,T. (2013) Spatial dynamics of chromosome translocations in living cells. *Science*, **341**, 660–664.
71. Rao,S.S., Huntley,M.H., Durand,N.C., Stamenova,E.K., Bochkov,I.D., Robinson,J.T., Sanborn,A.L., Machol,I., Omer,A.D., Lander,E.S. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.