Method

DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA

Daesik Kim[1,3] and Jin-Soo Kim[1,2]

[1]*Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea;* [2]*Center for Genome Engineering, Institute for Basic Science, Seoul 08826, Republic of Korea*

To investigate whether and how CRISPR-Cas9 on-target and off-target activities are affected by chromatin in eukaryotic cells, we first identified a series of identical endogenous DNA sequences present in both open and closed chromatin regions and then measured mutation frequencies at these sites in human cells using Cas9 complexed with matched or mismatched sgRNAs. Unlike matched sgRNAs, mismatched sgRNAs were highly sensitive to chromatin states, suggesting that off-target but not on-target DNA cleavage is hindered by chromatin. We next performed Digenome-seq using cell-free chromatin DNA (now termed DIG-seq) and histone-free genomic DNA in parallel and found that only a subset of sites, cleaved in histone-free DNA, were cut in chromatin DNA, suggesting that chromatin can inhibit Cas9 off-target effects in favor of its genome-wide specificity in cells.

[Supplemental material is available for this article.]

The type II CRISPR-Cas9 system, a form of adaptive immunity in eubacteria and archaea against foreign DNA elements, has been successfully repurposed for genome editing in higher eukaryotic cells (Cho et al. 2013; Cong et al. 2013; Mali et al. 2013), in which chromosomal DNA is wrapped around histones and packaged into chromatin. Although it has been shown that chromatin can impede Cas9–DNA interactions in vitro under cell-free conditions (Hinz et al. 2016; Horlbeck et al. 2016; Isaac et al. 2016), it remains unknown whether and how chromatin structure can affect Cas9 on-target and off-target activities in cells. Previous studies relied on statistical analyses of numerous single-guide RNAs (sgRNAs) that were targeted to different sequences and sites in the genome (Moreno-Mateos et al. 2015; Jensen et al. 2017), comparing apples with oranges, or reporter sequences randomly integrated at genomic sites (Chen et al. 2016; Daer et al. 2017; Kim et al. 2017), which can alter chromatin states. Wu et al. (2014) and Kuscu et al. (2014) independently showed, using chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq), that genome-wide off-target sites bound by catalytically deficient Cas9 (dCas9) are enriched in open chromatin regions, hinting at chromatin effects on off-target binding in cells. We and others found, however, that dCas9 off-target binding sites rarely overlap with Cas9 off-target cleavage sites (Tsai et al. 2015; Kim et al. 2016), calling for further studies.

In this study, we sought to investigate whether Cas9 nuclease activities in human cells can be affected by differential chromatin states, and to improve Digenome-seq, an in vitro method for profiling genome-wide CRISPR off-target sites using whole-genome sequencing (WGS) of digested genomic DNA, to account for chromatin states.
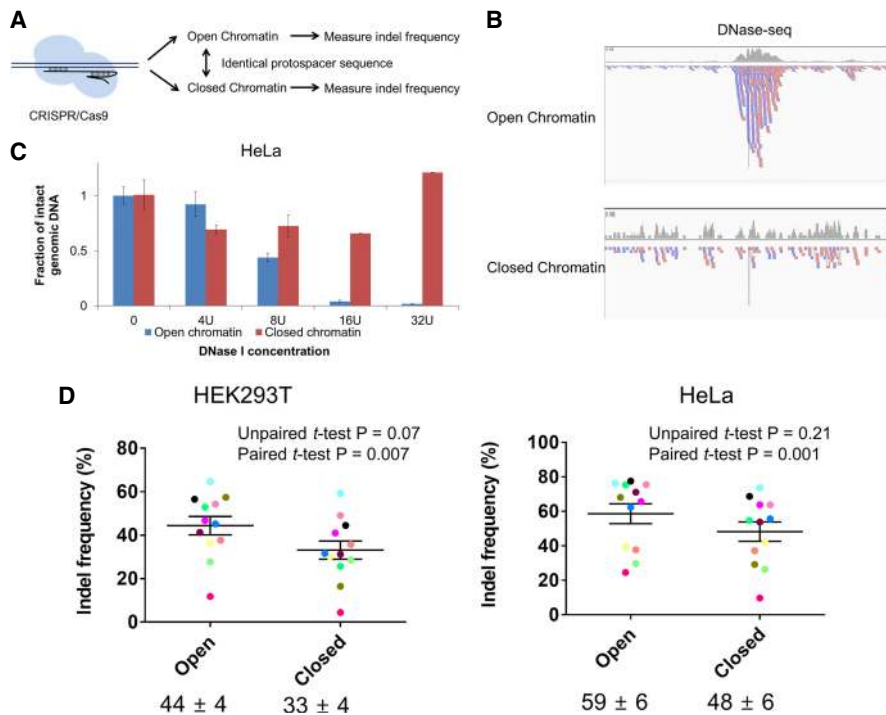
Results

Chromatin effects on CRISPR-Cas9 editing efficiency

We first identified 12 proto-spacer sequences found in both open and closed chromatin regions in the human genome, which can be targeted by *Streptococcus pyogenes* Cas9 (SpCas), using ENCODE DNase-seq data (Fig. 1A,B; The ENCODE Project Consortium. 2012) and performed DNase I digestion assays and quantitative PCR (qPCR) to confirm chromatin states in two widely different human cell lines, HeLa and HEK 293T. As expected, DNA sequences in open chromatin regions were more sensitive to DNase I digestion than those in closed chromatin regions (Fig. 1C; Supplemental Fig. 1). Next, we measured frequencies of small insertions and deletions (indels) induced, via error-prone nonhomologous end-joining (NHEJ) repair of DNA double-strand breaks (DSBs) at these 12 pairs of Cas9 target sites that were present in both open and closed chromatin regions (Fig. 1D; Supplemental Table 1). Because indel frequencies were highly variable among sgRNAs, there was no statistically significant difference between average indel frequencies in open chromatin regions and those in closed chromatin regions (unpaired *t*-test, $P = 0.07$ or 0.21 in HEK 293T or HeLa cells). A pair-wise comparison of each sgRNA, however, clearly showed that Cas9 induced indels with higher frequencies at sites in open chromatin regions than it did at respective sites with the same DNA sequence in closed chromatin regions (paired *t*-test, $P = 0.007$ or 0.001 in HEK 293T or HeLa cells, respectively) (Fig. 1D), suggesting that Cas9 cleaves target sites in open chromatin regions more efficiently than those in closed chromatin regions. On average, indel frequencies at open chromatin sites were higher than those at closed chromatin sites by a factor of 1.6 ± 0.2 or 1.3 ± 0.2 in HEK 293T or HeLa cells, respectively. These results show that on-target sites in open chromatin regions are marginally more accessible to Cas9 than are those in closed chromatin regions with the same DNA sequences.

[3]**Present address: Center for Genome Engineering, Institute for Basic Science, Seoul 08826, Republic of Korea**
Corresponding author: jskim01@snu.ac.kr

Figure 1. Effects of chromatin structure on Cas9 editing efficiency. (*A*) Schematic overview of the method for investigating the effects of chromatin structure on Cas9 editing efficiency. (*B*) Representative IGV images obtained using ENCODE DNase-seq data at two on-target sites with the same Cas9 target sequence present in both open and closed chromatin regions. (*C*) Relative fractions of intact genomic DNA not cleaved by DNase I measured using real-time quantitative PCR at the Cas9 target site in HeLa cells. The range 0–32U denotes the concentration of DNase I. Error bars, SEM from at least three independent experiments. (*D*) SpCas9-induced mutation frequencies at 12 pairs of endogenous target sites with the same DNA sequence present in both open and closed chromatin regions in HEK 293T or HeLa cells. Indel frequencies were measured using targeted deep sequencing. Pairs of Cas9 target sites are represented with dots with different colors. Mean indel frequencies ± SEM are shown (*n* = 12 target sites).

Fig. 3). Unlike on-target specific sgRNAs, off-target specific sgRNAs with 2-nt mismatches efficiently discriminated open chromatin sites from closed chromatin sites in both HeLa and HEK 293T cells. For example, two off-target–specific sgRNAs designed to match prevalidated off-target sites were quite active at the open chromatin site with an indel frequency of 3.4% or 1.3% in HeLa cells but were inactive at the closed chromatin site with an indel frequency of 0.05% or 0.03%, respectively, discriminating the two sites with the same DNA sequence by a factor of 68 or 43, respectively (Fig. 2B). In contrast, the on-target–specific sgRNA was almost equally efficient at the open chromatin site and the closed chromatin site with an indel frequency of 52% and 33%, respectively. We also tested off-target–specific sgRNAs at other target sites and obtained similar results in HEK 293T cells and HeLa cells (Supplemental Fig. 3). We conclude that Cas9 on-target interactions are strong enough to cause DNA cleavage and targeted mutagenesis at genomic sites regardless of the chromatin state, whereas Cas9 off-target interactions are in general much weaker and, thereby, are hindered by closed chromatin. This means that chromatin accessibility is in favor of genome-wide CRISPR-Cas9 specificity in human and other higher eukaryotic cells.

Chromatin effects on CRISPR-Cas9 off-target activity

Next, we investigated whether and how chromatin states affect Cas9 off-target effects using a series of mismatched sgRNAs. We chose two on-target sequences and measured indel frequencies at these sites in open and closed chromatin regions (Fig. 2A; Supplemental Fig. 2). The fully matched sgRNA specific to a proto-spacer sequence termed T1 was marginally more efficient at the open chromatin site than at the closed chromatin site by a factor of 1.3 (=44%/33%) in HEK 293T cells or 1.1 (=59%/48%) in HeLa cells. Interestingly, sgRNAs with 1- or 2-nucleotide (nt) mismatches were poorly active at the closed chromatin site but were still quite active at the open chromatin site. Thus, indel frequencies obtained with these mismatched sgRNAs at the open chromatin site were higher than at the closed chromatin sites by up to 530-fold in HEK 293T cells and 1100-fold in HeLa cells (Fig. 2A). We also tested a series of mismatched sgRNAs at the other on-target sequences present in open and closed chromatin regions and observed similar trends (Supplemental Fig. 2). These results show that mismatched sgRNAs, unlike matched sgRNAs, cannot efficiently guide Cas9 to sites in closed chromatin regions, and suggest that off-target effects are hindered by a closed chromatin structure.

To confirm that Cas9 access to off-target sites is more severely limited by chromatin states than Cas9 access to on-target sites, we tested sgRNAs matching prevalidated off-target sites at on-target sites in open and closed chromatin states (Fig. 2B; Supplemental

DIG-seq: Digenome-seq using native chromatin DNA

Having learned that chromatin is a barrier to Cas9 off-target DNA cleavage in cells, we sought to identify genome-wide Cas9 off-target sites in a chromatin context by using native chromatin DNA rather than histone-free genomic DNA in a digested genome sequencing (Digenome-seq) analysis (Kim et al. 2015). Native chromatin DNA isolated from HeLa and HEK 293T cells was incubated with the preassembled Cas9 protein (300 nM) and sgRNA (900 nM) ribonucleoprotein (RNP) complex for ~12 h and subjected to WGS (Fig. 3A; Kim et al. 2015). After aligning sequence reads to the human reference genome (hg19), an unusual pattern of vertical alignments, rather than staggered alignments, of sequence reads, caused by Cas9-catalyzed DNA cleavage, was observed using the Integrative Genomics Viewer (IGV) at the on-target site (Fig. 3B). We then identified genome-wide in vitro cleavage sites after assigning a DNA cleavage score, a measure of vertical alignments of sequence reads, to each nucleotide position in the entire human genome. With a cutoff score of 2.5 (Digenome v. 1.0) or 0.1 (Digenome v. 2.0) (Kim et al. 2016), we identified 15 and 44 in vitro cleavage sites, respectively, using Cas9 complexed to an sgRNA targeted to *HBB* (Fig. 3C,D; Supplemental Table 2). We obtained the same results when sequence reads were aligned to GRCh38 as to hg19 (Supplemental Table 3).

We also carried out DIG-seq using nuclei pellets rather than chromatin DNA. A vast majority (77%) of DIG-seq–positive sites identified using chromatin DNA were also cleaved using nuclei
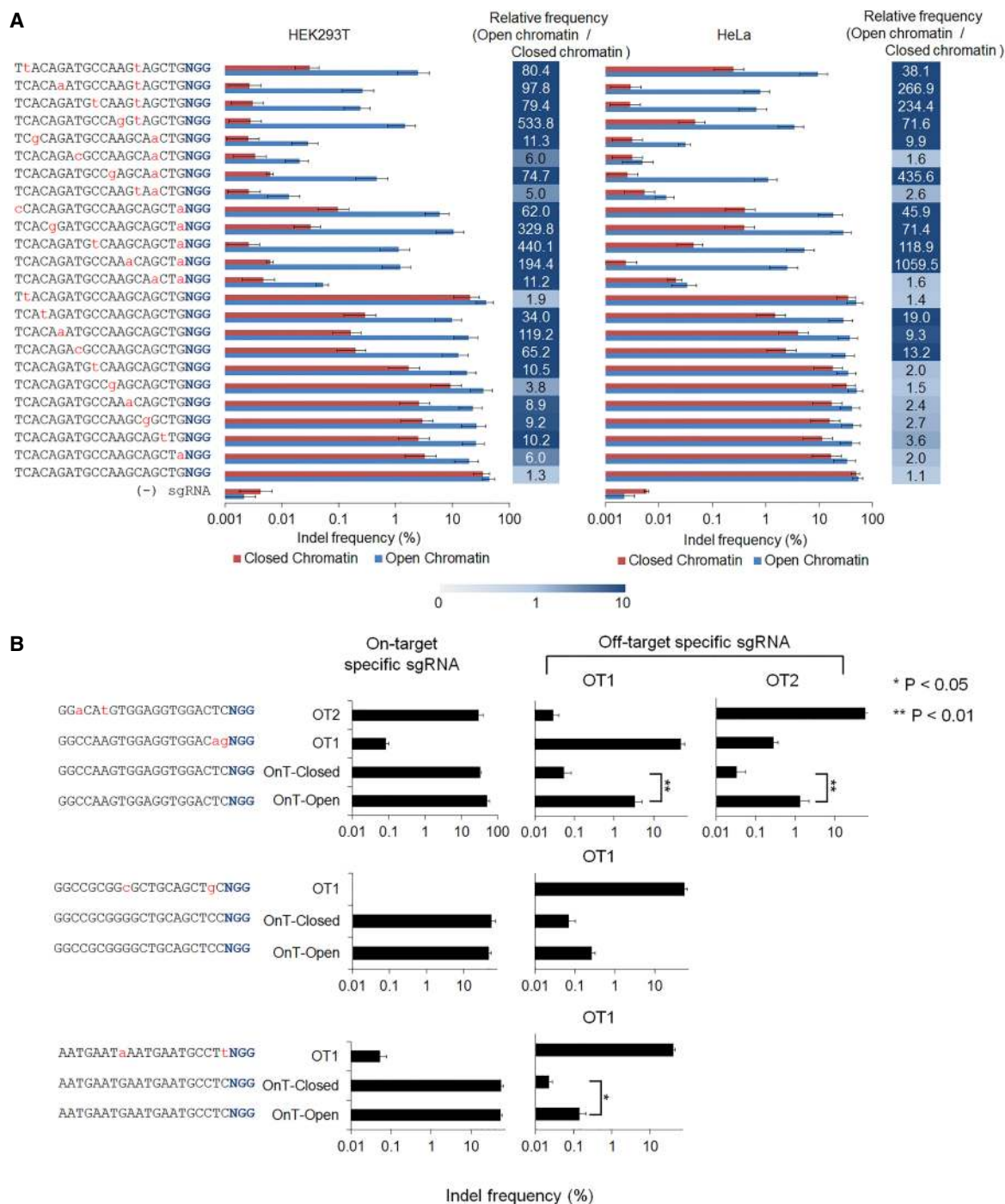
Figure 2. Effects of chromatin structure on Cas9 activity at off-target sites. (*A*) Mismatched sgRNAs that differed from the Cas9 target sequence by 1 or 2 nt were targeted to sites in both open (blue) and closed (red) chromatin regions in HEK 293T and HeLa cells. Indel frequencies were measured using targeted deep sequencing. Error bars, SEM ($n=3$). (*B*) Indel frequencies at the on-target sites in open and closed chromatin states using on-/off-target specific sgRNAs in HeLa cells. Indel frequencies were measured using targeted deep sequencing. Error bars, SEM ($n=3$).

pellets. Thus, we observed a strong correlation (Pearson's correlation coefficient $R^2 = 0.93$) between DNA cleavage scores obtained using nuclei pellets and those obtained using chromatin DNA (Supplemental Fig. 4).

We next performed DIG-seq using chromatin DNA isolated from K562 cells, T cells, and H9 human embryonic stem cells.

Most of the in vitro cleavage sites, especially those with high DNA cleavage scores, were commonly identified, regardless of cell type (Supplemental Table 2), suggesting that chromatin effects on Cas9 off-target DNA cleavage, at least with the sgRNA used in this study, were not much different in these cell lines.
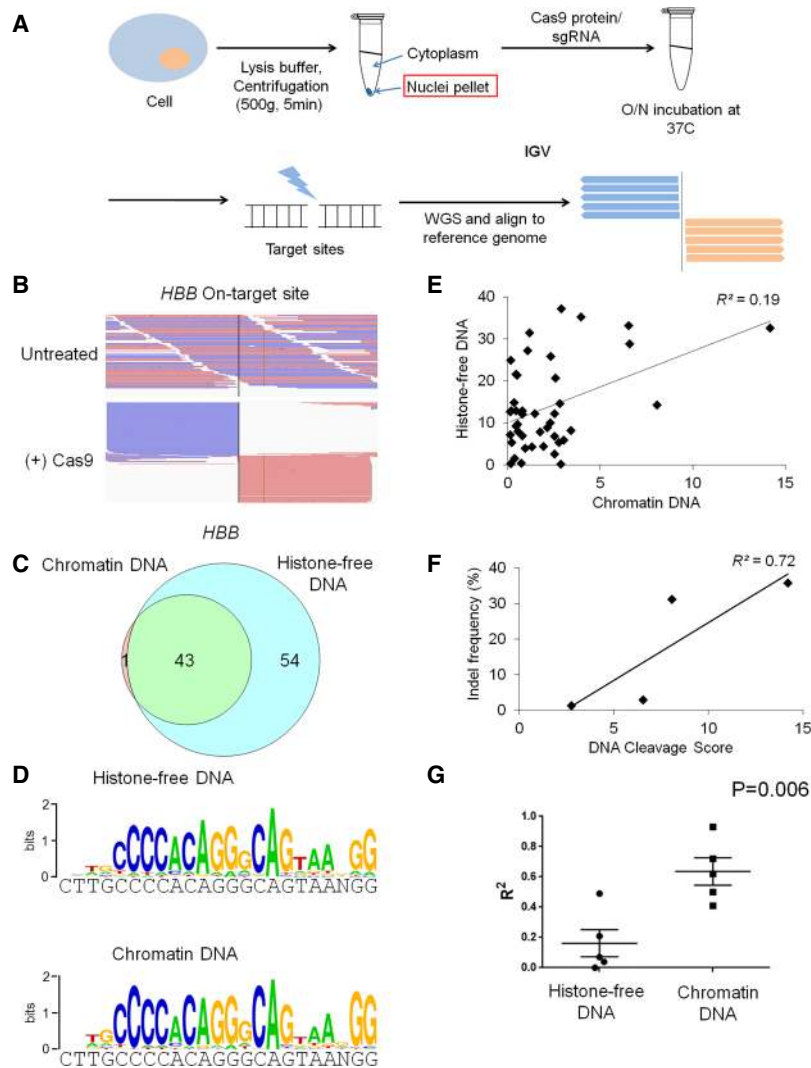
Figure 3. DIG-seq using native chromatin DNA. (*A*) Overview of DIG-seq to identify genome-wide Cas9 in vitro cleavage sites using native chromatin DNA. (*B*) A representative IGV image showing a staggered alignment (*top*) and a straight alignment (*bottom*) of whole-genome sequence reads at the on-target site. (*C*) A Venn diagram showing the number of in vitro cleavage sites identified by Digenome-seq using histone-free DNA or by DIG-seq using chromatin DNA with the *HBB*-targeted CRISPR-Cas9. (*D*) Sequence logos obtained via WebLogo using in vitro cleavage sites captured by Digenome-seq or DIG-seq. (*E*) Scatterplot of DNA cleavage scores at sites captured by Digenome-seq versus DIG-seq. (*F*) Scatterplot of indel frequencies versus DNA cleavage scores at sites captured by DIG-seq using HeLa chromatin DNA. (*G*) Pearson's correlation coefficients (*R*-square values) obtained with indel frequencies versus DNA cleavage scores at off-target sites identified by Digenome-seq using chromatin or histone-free DNA. Error bars, SEM ($n = 5$). The *P*-value was calculated by Student's *t*-test.

histone-free genomic DNA ($R^2 = 0.22$ [Digenome v. 1.0] or 0.19 [Digenome 2.0]) (Fig. 3E; Supplemental Fig. 5A).

Validating off-target effects at Digenome-identified sites

We next investigated the relationship between DNA cleavage scores and indel frequencies at off-target sites validated by next-generation sequencing (NGS). We measured indel frequencies at all of the 49 in vitro cleavage sites revealed with Digenome v. 1.0 in HeLa cells (Supplemental Fig. 5B; Supplemental Table 4) and identified four bona fide off-target sites at which indel frequencies were greater than background sequencing error rates. No indels were detectably induced at 34 in vitro cleavage sites detected by Digenome-seq using histone-free DNA but missed by DIG-seq using chromatin DNA (Supplemental Fig. 5B; Supplemental Table 5). There was essentially no correlation between DNA cleavage scores obtained using histone-free DNA and indel frequencies at these validated off-target sites ($R^2 = 0.07$) (Supplemental Fig. 5C). Interestingly, however, we observed a strong correlation between DNA cleavage scores obtained using chromatin DNA and indel frequencies ($R^2 = 0.72$) (Fig. 3F).

We next performed DIG-seq with seven additional sgRNAs individually (Supplemental Table 4) and compared the results with those obtained using histone-free DNA. As expected, fewer sites were cleaved in chromatin DNA than in histone-free DNA by Cas9. Thus, a total of eight sgRNAs in complex with Cas9 cleaved chromatin DNA and histone-free DNA at 18 ± 6 and 85 ± 19 sites, on average, respectively. Most sites (66%–100%, on average 91%) cleaved in chromatin DNA were also cleaved in histone-free DNA (Fig. 3C; Supplemental Fig. 6; Supplemental Table 6).

We measured indel frequencies at these in vitro cleavage sites identified by DIG-seq using chromatin DNA to validate Cas9 off-target effects in cells. Forty-nine out of 138 sites tested in this study were validated to be bona fide off-target sites. Indel frequencies at these validated off-target sites ranged from 0.07% (*FANCF*-5 in Supplemental Table 4) to 39% (*EMX*1-4 in Supplemental Table 4). Validation rates of DIG-seq were higher than those of Digenome-seq. Thus, among 15 DIG-seq–positive sites identified using the *HBB* sgRNA via Digenome v. 1.0, four sites were validated using deep sequencing, resulting in a validation rate of 27%, whereas among 48 Digenome-seq–positive sites via Digenome v. 1.0, the same four sites were validated, resulting in a validation rate of 8.3%. Likewise, the DIG-seq validation rate was 100% (=1/1) using the *RNF2* sgRNA, whereas the Digenome-

We compared in vitro cleavage sites obtained using native chromatin DNA with those obtained using histone-free DNA. Use of histone-free DNA in a Digenome-seq analysis yielded many additional in vitro cleavage sites. Thus, with Digenome v. 1.0, 15 or 48 sites were cleaved using native chromatin DNA and histone-free DNA, respectively. Fourteen out of 15 (=93%) sites captured by DIG-seq were also identified by Digenome-seq using histone-free DNA. With Digenome v. 2.0, 44 or 97 sites were cleaved using native chromatin DNA and histone-free DNA, respectively. Forty-three out of 44 (=98%) sites captured using chromatin DNA were also identified using histone-free DNA. We found, however, that DNA cleavage scores at these in vitro cleavage sites obtained using native chromatin were poorly correlated with those obtained using

seq validation rate was 23% (=3/13). The sites detected exclusively by DIG-seq (one site for the *HBB* sgRNA, 10 sites for the *FANCF* sgRNA, and one site for the *EMX1* sgRNA) were not validated by targeted deep sequencing. It is possible that these sites were mutated at frequencies below detection limits of targeted amplicon sequencing. This means that DIG-seq using chromatin DNA is as sensitive and comprehensive as Digenome-seq using histone-free DNA. Again, we observed a strong correlation between DNA cleavage scores obtained using chromatin DNA and indel frequencies determined using NGS at validated off-target sites (Fig. 3G; Supplemental Fig. 7). In contrast, DNA cleavage scores obtained using histone-free DNA were poorly correlated with indel frequencies. This result suggests that a DNA cleavage score obtained using chromatin DNA rather than histone-free DNA could serve as an estimate of genome editing efficiency in cells. It is often impractical to validate off-target effects in cells using NGS at all the candidate sites cleaved in vitro because the number of such sites can reach hundreds to thousands, depending on the sgRNAs (Tsai et al. 2015; Kim et al. 2016). In these cases, one could choose the sites with the top cleavage scores obtained using chromatin DNA for validating off-target effects via NGS.

We also found DIG-seq–positive sites were more likely to fall within open-chromatin regions than were Digenome-seq–positive sites. Among 139 DIG-seq–positive sites identified using the other seven sgRNAs, seven sites (=5.0%) were found in DNase I hypersensitive regions, whereas among 358 Digenome-seq–positive sites identified using the same seven sgRNAs, just four sites (=1.1%) were found in DNase I hypersensitive regions (Supplemental Tables 4, 7).

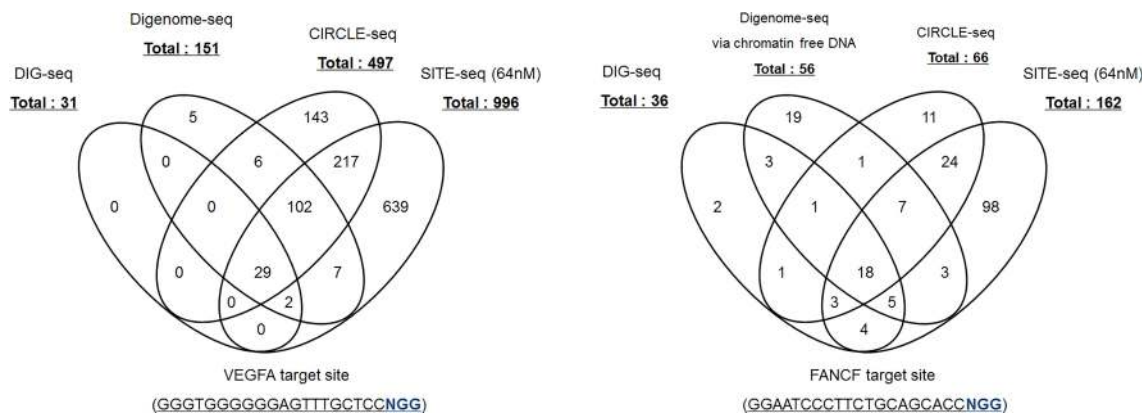### Digenome-seq and other methods

We also compared our new method with CIRCLE-seq (Tsai et al. 2017) and SITE-seq (Cameron et al. 2017), recently reported cell-free methods for characterizing CRISPR off-target effects in vitro. Among the eight sgRNAs we analyzed in this study, two sgRNAs, namely those targeted to the *VEGFA* site and the *FANCF* site, had also been tested by Tsai et al. (2017) and Cameron et al. (2017), although they used genomic DNA isolated from different cell lines. We found that none of these methods, including Digenome-seq with histone-free DNA, were comprehensive (Fig. 4). SITE-seq produced the most outputs: 996 sites and 162 sites were identified using the *VEGFA*-specific sgRNA and the *FANCF*-specific sgRNA, respectively (Cameron et al. 2017). DIG-seq yield-

ed by far the fewest sites, 31 and 36, respectively, even with the more comprehensive Digenome v. 2.0 program (Kim et al. 2016). Among these sites, 29 out of 31 (94%) and 21 out of 36 (58%), respectively, were also captured by both CIRCLE-seq and SITE-seq. We validated off-target effects in cells at the 29 sites commonly identified via the four different methods using the *VEGFA*-specific sgRNA and found that 16 sites were bona fide off-targets, resulting in a validation rate of 62% (=18/29). Because CIRCLE-seq and SITE-seq, in general, yield more outputs than Digenome-seq using histone-free DNA, validation rates with sites identified using these cell-free methods are much lower, ranging from 10% (SITE-seq) (Cameron et al. 2017) to 29% (CIRCLE-seq) (Tsai et al. 2017).

Discussion

Genome-wide CRISPR off-target sites can be identified by a number of unbiased cell-free or cell-based methods (Koo et al. 2015; Tsai and Joung 2016). Each of these methods has its own pros and cons. DIG-seq is unique in that it is an in vitro (cell-free) method but nevertheless accounts for chromatin structure in eukaryotic cells. Chromatin DNA can be easily prepared from any cells of interest by lysis and centrifugation and used for digestion with Cas9 or other nucleases in vitro. In contrast, GUIDE-seq, a widely used cell-based method, requires transfection of double-stranded oligonucleotides into cells (Tsai et al. 2015). GUIDE-seq cannot be used in certain cells that are refractory to transfection. Furthermore, double-stranded oligonucleotides are cytotoxic to many primary cells. DIG-seq does not rely on transfection and is thereby free from these limitations. In addition, DIG-seq retains all of the advantages of Digenome-seq over other cell-free or cell-based methods, which include no adapter ligation or DNA amplification prior to deep sequencing, no need for homology searches (Digenome v. 1.0) to identify potential off-target sites, and no DNA resection in vitro by cellular repair enzymes, making it possible to pinpoint DSB ends precisely.

Cell-free methods provide a list of potential off-target sites that are cleaved in vitro. No mutations are, however, detectably induced at most of these sites in cells (Kim et al. 2015, 2016). According to our Digenome analysis using chromatin DNA, most of these in vitro cleavage sites are not cleaved in a chromatin context. NGS validation is an essential step for confirming and measuring genome-wide off-target effects but is time-consuming, expensive, and labor-intensive. In this regard, DIG-seq is advantageous over

Figure 4. Comparison of DIG-seq with other in vitro methods. Venn diagrams showing the number of in vitro cleavage sites identified by DIG-seq, Digenome-seq, CIRCLE-seq, and SITE-seq.

other cell-free methods because it provides a list of much fewer sites that are cleaved in a chromatin context than do other methods.

We note that there were several off-target candidate sites (e.g., HBB-3, HBB-5, HBB-6, etc., in Supplemental Table 4) captured by DIG-seq with relatively high DNA cleavage scores that were not validated by targeted deep sequencing. It is possible that these sites were mutated in cells at frequencies blow detection limits of deep sequencing, which range from 0.001%–1% (0.1% on average). Another possibility is that these sites were cleaved in cells efficiently but repaired seamlessly either via error-free NHEJ or HDR.

In conclusion, we found that CRISPR-Cas9 access to on-target and off-target sites is limited by chromatin structure in cells and that Cas9 interactions with sgRNA-mismatched sites are much more sensitive to the chromatin state than are those with matched sites, suggesting that chromatin enhances genome-wide CRISPR specificity in human and other eukaryotic cells. These results prompted us to develop DIG-seq using chromatin DNA rather than histone-free DNA in vitro for characterizing genome-wide CRISPR off-target effects in a chromatin context. We expect that DIG-seq will be widely used for identifying CRISPR-Cas9 and other nuclease off-target sites sensitively and comprehensively to facilitate therapeutic genome editing.

Methods

Cas9 protein purification and in vitro sgRNA transcription

Recombinant Cas9 protein was purchased from ToolGen. sgRNAs were synthesized by in vitro transcription using T7 RNA polymerase as described previously (Kim et al. 2016). Briefly, sgRNA templates were incubated with T7 RNA polymerase in reaction buffer (40 mM Tris-HCl, 6 mM $MgCl_2$, 10 mM DTT, 10 mM NaCl, 2 mM spermidine, NTPs, and RNase inhibitor, at pH 7.9) for 8 h at 37°C and mixed with DNase I to remove sgRNA template DNA. Transcribed sgRNAs were purified using PCR purification kits (Macrogen).

Cell culture and transfection conditions

HEK 293T cells (ATCC CRL-11268) and HeLa cells (ATCC CCL-2) were cultured in DMEM media supplemented with 10% FBS and 1% penicillin/streptomycin (Welgene). HEK 293T cells (1.5×10^5) or HeLa cells (8×10^4) were seeded on 24-well plates and co-transfected with the Cas9 expression plasmid (500 ng) and the sgRNA-encoding plasmid (500 ng) using Lipofectamine 2000 (Life Technologies). K562 cells (ATCC) were maintained in RPMI medium supplemented with 10% FBS and 1% penicillin/streptomycin (Welgene).

DNase I digestion assays and qPCR

We washed 2×10^7 cells twice with PBS and incubated them with DNase I (4–32 units) in a reaction volume of 120 µL (10 mM NaCl, 10 mM Tris-HCl, 3 mM $MgCl_2$, and 0.1% NP-40, at pH 7.4) for 30 min at 37°C. One hundred sixty microliters of 50 mM EDTA was added to stop the reaction. DNase I–digested DNA was purified with a DNeasy tissue kit (Qiagen). Digested DNA was mixed with KAPA SYBR FAST qPCR master mix (Kapa Biosystems) and analyzed by real-time qPCR. The fraction of intact genomic DNA was measured using the comparative C_T method; $2^{-\Delta\Delta CT} =$ [(C_T site of interest of DNase I untreated or treated sample – C_T internal control of DNase I untreated or treated sample) – (C_T site of interest of DNase I untreated sample – C_T internal control of DNase I untreated sample)] (Schmittgen and Livak 2008).

Identification of identical Cas9 target sequences that are present in both open and closed chromatin regions

To find Cas9 target sequences present in both open and closed chromatin regions, we first identified Cas9 target sequences that occur twice in the genome. By using ENCODE DNase-seq data obtained from HEK293T and HeLa cells, we manually identified Cas9 target sequences with a high sequencing depth at one site and a low depth at the other site, corresponding to an open chromatin site and a closed chromatin site, respectively.

In vitro cleavage of native chromatin DNA

We lysed 5×10^5 cells with lysis buffer (1× PBS, 0.4% NP-40, and 3 mM $MgCl_2$) and centrifuged them at 500g for 5 min. After removal of supernatant, nuclei pellets were mixed with Nuc-lysis solution (10 mM EDTA, 0.5 mM EGTA, 0.1% Triton X-100) and centrifuged at 500g for 5 min. Nuclei pellets or native chromatin were incubated with 300 nM of Cas9 and 900 nM of sgRNA for 8 h in reaction buffer (100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, and 100 µg/mL BSA, at pH 7.9). Digested genomic DNA was treated with RNase A (50 µg/mL) to degrade sgRNAs and was purified again with a DNeasy tissue kit (Qiagen).

WGS and Digenome sequencing

One microgram of genomic DNA was sonicated to 400- to 500-bp fragments using the Covaris system (Life Technologies), and overhangs were removed using the end repair mix. Fragmented DNA was ligated with adapters to produce libraries, which were then subjected to WGS using an Illumina HiSeq X Ten Sequencer at Macrogen. WGS was performed at a sequencing depth of 30–40× with the following mapping program and parameters: Isaac aligner, ver. 01.14.03.12; human genome reference, hg19 or GRCh38; base quality cutoff, 15; keep duplicate reads, yes; variable read length support, yes; realign gaps, no; and adapter clipping, yes (adapter: AGATCGGAAGAGC*,*GCTCTTCCGATCT). DNA cleavage scores at the DNA target sites were calculated using the Digenome 1.0 and Digenome 2.0 programs (https://github.com/chizksh/digenome-toolkit2). As described in our previous paper (Kim et al. 2017), the cutoff value was determined experimentally. Briefly, we counted the number of sites whose DNA cleavage scores were over a cutoff value that ranged from 0.0001 to 10 and the number of PAM-containing sites with 10 or fewer mismatches among the sites with scores over the cutoff value (Supplemental Fig. 8). We chose a cutoff value of 0.1 because WGS data obtained using intact genomic DNA, which served as a negative control, did not yield any false-positive sites with this cutoff score (Digenome v. 2.0).

Targeted deep sequencing

Genomic DNA segments spanning the on-target and potential off-target sites were amplified by Phusion polymerase (New England Biolabs), and PCR amplicons were subjected to paired-end sequencing using Illumina MiSeq. The resulting sequencing files were subjected to Cas-Analyzer (http://www.rgenome.net/cas-analyzer/) to calculate indel frequencies (Park et al. 2017).

Data access

The next-generation sequencing data from this study have been submitted to the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) under accession numbers SRP067307 and SRP158339.

References

Cameron P, Fuller CK, Donohoue PD, Jones BN, Thompson MS, Carter MM, Gradia S, Vidal B, Garner E, Slorach EM, et al. 2017. Mapping the genomic landscape of CRISPR–Cas9 cleavage. *Nat Methods* **14:** 600–606. doi:10.1038/nmeth.4284

Chen X, Rinsma M, Janssen JM, Liu J, Maggio I, Gonçalves MA. 2016. Probing the impact of chromatin conformation on genome editing tools. *Nucleic Acids Res* **44:** 6482–6492. doi:10.1093/nar/gkw524

Cho SW, Kim S, Kim JM, Kim JS. 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31:** 230–232. doi:10.1038/nbt.2507

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339:** 819–823. doi:10.1126/science.1231143

Daer RM, Cutts JP, Brafman DA, Haynes KA. 2017. The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. *ACS Synth Biol* **6:** 428–438. doi:10.1021/acssynbio.5b00299

The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489:** 57–74. doi:10.1038/nature11247

Hinz JM, Laughery MF, Wyrick JJ. 2016. Nucleosomes selectively inhibit Cas9 off-target activity at a site located at the nucleosome edge. *J Biol Chem* **291:** 24851–24856. doi:10.1074/jbc.C116.758706

Horlbeck MA, Witkowsky LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, Weissman JS. 2016. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *eLife* **5:** e12677. doi:10.7554/eLife.12677

Isaac RS, Jiang F, Doudna JA, Lim WA, Narlikar GJ, Almeida R. 2016. Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *eLife* **5:** e13450. doi:10.7554/eLife.13450

Jensen KT, Fløe L, Petersen TS, Huang J, Xu F, Bolund L, Luo Y, Lin L. 2017. Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett.* **591:** 1892–1901. doi:10.1002/1873-3468.12707

Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim JI, Kim JS. 2015. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* **12:** 237–243. doi:10.1038/nmeth.3284

Kim D, Kim S, Kim S, Park J, Kim JS. 2016. Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res* **26:** 406–415. doi:10.1101/gr.199588.115

Kim HK, Song M, Lee J, Menon AV, Jung S, Kang YM, Choi JW, Woo E, Koh HC, Nam JW, et al. 2017. *In vivo* high-throughput profiling of CRISPR–Cpf1 activity. *Nat Methods* **14:** 153–159. doi:10.1038/nmeth.4104

Koo T, Lee J, Kim JS. 2015. Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9. *Mol Cells* **38:** 475–481. doi:10.14348/molcells.2015.0103

Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. 2014. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Boitechnol* **32:** 677–683. doi:10.1038/nbt.2916

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* **339:** 823–826. doi:10.1126/science.1232033

Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK, Giraldez AJ. 2015. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting *in vivo*. *Nat Methods* **12:** 982–988. doi:10.1038/nmeth.3543

Park J, Lim K, Kim JS, Bae S. 2017. Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics* **33:** 286–288. doi:10.1093/bioinformatics/btw561

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C_T method. *Nat Protoc* **3:** 1101–1108.

Tsai SQ, Joung JK. 2016. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat Rev Genet* **17:** 300–312. doi:10.1038/nrg.2016.28

Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, Iafrate AJ, Le LP, et al. 2015. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33:** 187–197. doi:10.1038/nbt.3117

Tsai SQ, Nguyen NT, Malagon-Lopez J, Topkar VV, Aryee MJ, Joung JK. 2017. CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nat Methods* **14:** 607–614. doi:10.1038/nmeth.4278

Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, Cheng AW, Trevino AE, Konermann S, Chen S, et al. 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* **32:** 670–676. doi:10.1038/nbt.2889

DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA

Daesik Kim and Jin-Soo Kim

Supplemental Material	http://genome.cshlp.org/content/suppl/2018/11/09/gr.236620.118.DC1
References	This article cites 23 articles, 4 of which can be accessed free at: http://genome.cshlp.org/content/28/12/1894.full.html#ref-list-1
Open Access	Freely available online through the *Genome Research* Open Access option.
Creative Commons License	This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions