Published in final edited form as:

Nat Protoc. 2018 November ; 13(11): 2615–2642. doi:10.1038/s41596-018-0055-0.

Defining CRISPR-Cas9 genome-wide nuclease activities with CIRCLE-seq

Cicera R. Lazzarotto¹, Nhu T. Nguyen^{2,5}, Xing Tang¹, Jose Malagon-Lopez^{2,3,4}, Jimmy Guo², Martin J. Aryee^{2,3,4}, J. Keith Joung^{2,3}, and Shengdar Q. Tsai¹

¹Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN, USA

²Molecular Pathology Unit, Massachusetts General Hospital, Charlestown, MA, USA

³Department of Pathology, Harvard Medical School, Boston, MA, USA

⁴Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

⁵Present address: Cutaneous Biology Research Center, Department of Dermatology, Massachusetts General Hospital, and Harvard Medical School, Boston, MA USA

Abstract

Circularization for *in vitro* reporting of cleavage effects by sequencing (CIRCLE-seq) is a sensitive and unbiased method for defining the genome-wide activity (on-target and off-target) of CRISPR-Cas9 nucleases by selective sequencing of nuclease-cleaved genomic DNA. Here we describe a detailed experimental and analytical protocol for CIRCLE-seq. The principle of our method is to generate a library of circularized genomic DNA with minimized numbers of free ends. Highly purified genomic DNA circles are treated with CRISPR-Cas9 ribonucleoprotein complexes, and nuclease-linearized DNA fragments are then ligated to adapters for high-throughput sequencing. Primary advantages of CIRCLE-seq when compared to other *in vitro* methods for defining genome-wide genome editing activity include: 1) high enrichment for sequencing nuclease-cleaved genomic DNA/low background enabling sensitive detection with low sequencing depth requirements and 2) paired-end reads can contain complete information on individual nuclease cleavage sites enabling use in species without high-quality reference genomes. The entire protocol can be completed in 2 weeks, with 3 days for library preparation.

Keywords

CRISPR-Cas; off-target activity; Cas9; off-target effect; genome editing; CRISPR; CIRCLE-seq

Correspondence should be addressed to: shengdar.tsai@stjude.org.

Author Contributions

C.R.L. and S.Q.T. wrote the manuscript with input from all authors. N.T.N. and S.Q.T. developed the original experimental protocol in the lab of J.K.J. X.T., J.M-L., M.J.A., and S.Q.T. contributed to CIRCLE-seq software analysis pipeline. C.R.L. performed experiments and data analysis.

Competing Financial Interests

J.K.J. has financial interests in Beam Therapeutics, Editas Medicine, Monitor Biotechnologies, Pairwise Plants, Poseida Therapeutics, and Transposagen Biopharmaceuticals. S.Q.T. has financial interests in Monitor Biotechnologies. J.K.J.'s interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies. J.K.J. and S.Q.T. are co-inventors on a patent describing the CIRCLE-seq method that has been licensed to Monitor Biotechnologies.

Tweet: Defining CRISPR-Cas nuclease on- and off-target activities genome-wide with CIRCLE-seq

INTRODUCTION

Targeted genome editing is enabled by engineered nucleases such as meganucleases¹, zinc fingers (ZFNs)², transcription activator-like effector nucleases (TALENs)^{3,4} and CRISPR-Cas nucleases^{4,5}, enzymes that can introduce targeted DNA double-stranded breaks (DSBs) into the genome of living cells. Nuclease-induced DSBs can typically undergo repair by one of the two major DNA damage repair pathways: error-prone non-homologous end-joining (NHEJ) often resulting in variable-length insertion or deletion (indel) mutations, or relatively precise homology directed repair (HDR)^{4,6}

Engineered nucleases can introduce undesired off-target mutations in living cells and organisms^{7,8}, potentially leading to unintended effects in research and clinical applications. Unlike integrating viral methods for genome modification, nuclease-induced indel mutations cannot be easily detected by anchored PCR amplification of an expected sequence. Therefore, sensitive and unbiased genome-wide methods are required to fully understand the genomic locations and frequency of nuclease-induced DSB activity. While there are many situations where it may be helpful to appreciate the genome-wide activities of specific genome editing nucleases, it is particularly important in the context of clinical applications where millions or billions of cells may be modified, and rare off-target mutations could confer an undesired proliferative advantage and increased oncogenic potential.

Development of CIRCLE-seq

Here we present a detailed working protocol for identifying the genome-wide activity of CRISPR-Cas nucleases, called Circularization for *in vitro* Reporting of Cleavage Effects by sequencing or CIRCLE-seq⁹. CIRCLE-seq is an *in vitro* screening strategy for defining nuclease-induced off-target DSBs in a sensitive and unbiased fashion that overcomes sensitivity limitations of cell-based methods for genome-wide off-target detection. CIRCLE-seq is based on the circularization of randomly sheared genomic DNA using a restriction-enzyme independent strategy followed by exonuclease treatment, thereby generating a population of covalently closed circular double-stranded DNA molecules. Subsequently, nuclease-induced cleavage of these covalently closed DNA molecules at on- and off-target sites releases free DNA ends that serve as substrates for adapter ligation, enabling the selective sequencing of Cas9-cleaved genomic DNA molecules. Each pair of reads from Cas9-cleaved DNA can contain complete sequence information for a single on-target or off-target site. Because CIRCLE-seq reduces the sequencing coverage requirements for sensitive detection of CRISPR-Cas nuclease cleavage activity from human genomic DNA, the method can be run in any lab with access to high-throughput sequencing capability.

Comparison with other methods

There are a range of genome-wide methods that have recently been developed for detecting off-target activity of engineered nucleases¹⁰ (Table 1). These methods can be divided into two broad categories: cell-based and cell-free (*in vitro*) approaches. In general, cell-based approaches can directly discover off-target sites that are cleaved in a specific cell type; however, methods also require the ability to culture and transfect cells of interest and are

subject to potential fitness effects conferred by off-target mutations. By contrast, *in vitro* methods can be more highly sensitive and more scalable than cell-based methods. Both approaches ultimately require follow-up cell-based validation to confirm whether off-target sites identified by these methods result in *bona fide* off-target mutagenesis in the actual cells or tissues modified by genome-editing nucleases.

Integration-defective lentiviral (IDLV)¹¹ capture and genome-wide unbiased identification of DSBs enabled by sequencing (GUIDE-seq)^{12,13} are cell-based, end-capture methods for defining sites of nuclease-induced DSBs by incorporation of exogenous DNA sequences followed by tag-specific amplification. IdLv capture was the first method for unbiased identification of nuclease-induced DSBs¹⁴; its sensitivity is limited to sites with mutagenesis frequencies of approximately 5%. GUIDE-seq is based on the principle of efficient incorporation of a short, end-protected double-stranded DNA tag into the sites of nuclease-induced DSBs, tag-specific amplification and high-throughput sequencing; it has improved sensitivity compared with IDLV capture and can detect off-target sites that occur with frequencies as low as ~0.1%.

High-throughput genome-wide translocation (HTGTS)^{15,16} is a cell-based technique for off-target discovery that relies on the detection of translocations between two DSBs, a nuclease-induced 'bait' DSB and an off-target 'prey' DSBs. Translocations occur more frequently with sites on the same chromosome or chromosomes that are in close nuclear proximity. Thus, translocation frequency is not predicted to correlate well with nuclease-induced mutation frequency.

BLESS (breaks labeling, enrichment on streptavidin and next-generation sequencing)¹⁷, BLISS (breaks labeling *in situ* and sequencing)¹⁸, and DSBCapture¹⁹ are *in situ* methods for detecting genome-wide nuclease-induced DSBs in fixed cells. BLESS, BLISS and DSBCapture detect nuclease activity by ligation of adapters to unrepaired DSBs in fixed cells and do not depend on (1) the delivery and incorporation of exogenous DNA and on (2) the endogenous cellular DNA damage repair machinery for detection of cuts in the DNA. End-seq²⁰ is a similar *in situ* method for detecting genome-wide nuclease activity that detects DSBs in cells embedded in agarose rather than fixed cells. One limitation of these *in situ* methods is that they only capture a snapshot of DSBs that exist at the moment in time in which the cells are fixed and are insensitive to previously-repaired DSBs.

Diggenome-seq^{21,22} is an *in vitro* and PCR-free method for detecting Cas9 cut sites in bulk genomic DNA by whole-genome sequencing (WGS). This method is performed by digesting genomic DNA with ribonucleoprotein (RNP) complex *in vitro*, random DNA shearing, adapter ligation to all free ends (nuclease- and non-nuclease-induced) and WGS. Genomic DNA fragments are sequenced to approximately 30-50× coverage using ~400 million paired end reads, and mapped to a reference genome, where positions of uniformly mapping ends are computationally identified as sites of likely Cas9 DSBs. Due to the high uniform background of random genomic DNA reads, it can be challenging to distinguish low-frequency nuclease-induced cleavage events from background.

SITE-seq²³ is another *in vitro* method for enrichment and sequencing of Cas9-cleaved genomic DNA, based on the isolation of high-molecular weight genomic DNA to minimize the number of available DNA ends, cleavage with Cas9 RNP, biotinylated adapter ligation, enzymatic fragmentation, a second adapter ligation step to the enzymatically fragmented DNA ends and enrichment. SITE-seq and Digenome-seq reads yield only one-half of the cleavage site, in contrast to CIRCLE-seq reads which contain both halves.

Advantages

CIRCLE-seq has a number of advantages over other methods. When compared with Digenome-seq, CIRCLE-seq virtually eliminates the high background of random reads. Due to high levels of enrichment for nuclease-cleaved genomic DNA fragments, fairly minimal amounts of sequencing reads are required to successfully perform the assay (~3-5 million).

CIRCLE-seq is more sensitive than cell-based methods for genome-wide off-target detection. For most targets, CIRCLE-seq can identify all off-target sites in human gDNA found by GUIDE-seq, one of the most sensitive cell-based methods. In addition, CIRCLE-seq also identified new *bona fide* off-target sites that occur in human cells, demonstrating that it can detect novel sites that are below the limits of detection of GUIDE-seq¹².

A unique feature of the CIRCLE-seq method is that genomic DNA circles linearized by CRISPR-Cas nuclease cleavage contain both ends of the cleaved DNA molecule. Thus, CIRCLE-seq enables reference-genome independent off-target discovery, an advantage in organisms lacking full genomic sequence or those with high genetic variability.

Because CIRCLE-seq is a cell-free technique that directly maps nuclease-induced DSBs, it does not rely on cellular DNA damage repair pathways for detection and is not influenced by Cas9 delivery method.

Limitations

CIRCLE-seq may detect some sites that cannot be confirmed in cells because: 1) some sites that can be cleaved in purified genomic DNA *in vitro* may not be cleaved in cells due to limited chromatin accessibility or other cellular factors; and 2) some sites may fall below the limits of detection current high-throughput sequencing technologies. Conversely, typically all sites that are detected by cell-based approaches will be captured by *in vitro* techniques like CIRCLE-seq.

Each CIRCLE-seq sample requires a relatively large amount of genomic DNA (gDNA) for circularization (approximately 25 µg), which can be a limiting factor depending on the availability of the cellular source of genomic DNA. This represents the yield of genomic DNA from approximately 5 million diploid cells.

Choice of experimental method

The choice of method for determining genome-wide off-target activity of CRISPR-Cas nucleases depends on the primary experimental question at hand. *In vitro* methods such as CIRCLE-seq, Digenome-seq, and SITE-seq are the most comprehensive, though activity in cells will need to be verified in subsequent sequencing experiments. End-capture methods,

like GUIDE-seq, directly and quantitatively measure off-target mutagenesis in cells, but have limitations for consistently detecting sites with mutations with frequencies of lower than 0.1%. *In situ* ligation-based techniques detect unrepaired DSBs and may be useful for understanding the kinetics of on-target and off-target genome editing.

EXPERIMENTAL DESIGN

Overview of workflow

The CIRCLE-seq protocol can be divided into eight stages (See Fig. 1): (i) cell culture or tissue harvesting, (ii) genomic DNA extraction, (iii) sgRNA cloning and *in vitro* transcription, (iv) *in vitro* cleavage test of RNP complex in a PCR amplicon, (v) genomic DNA shearing, (vi) CIRCLE-seq library preparation for intramolecular circularization and enzymatic selection of circularized gDNA, (vii) circularized gDNA RNP *in vitro* cleavage, (viii) Next-Generation Sequencing (NGS) library preparation, NGS and data analysis.

gDNA input

CIRCLE-seq requires approximately 25 µg of genomic DNA (input amount before shearing) per nuclease-treated or control sample. The expected yield of purified, circularized DNA is approximately 1-2%.

sgRNA

sgRNA can be obtained by *in vitro* transcription from a plasmid template or, alternatively, by commercial synthesis. It is ideal to use the same source of sgRNA as used in cell-based experiments.

Controls required

One negative control is required for each unique source of genomic DNA used in CIRCLE-seq experiments and should be accounted for when considering gDNA input requirements. The negative control is processed in exactly the same manner as other samples, but with water substituted for CRISPR-Cas RNP at the nuclease cleavage step. This required negative control sample is sequenced along with all nuclease-treated samples.

Reaction master mixes

Include a standard excess of 15-20% when making all library preparation reaction master mixes to account for losses while pipetting.

Genomic DNA shearing and gDNA circularization

Purified genomic DNA is sheared to an average size of 300 bp. The fragmented DNA is ligated (by T4 DNA ligase) to uracil-containing stem-loop adapter via a standard high-throughput sequencing library preparation method²⁴. Adapter ligated DNA molecules are selected over the molecules that do not have adapters ligated in both ends (using Lambda exonuclease and *E. coli* Exonuclease I). Adapter-ligated DNA is then enzymatically treated (by USER enzyme and T4 PNK) to expose palindromic 4-nt overhangs at both ends. DNA circularization (by T4 DNA ligase) is performed at low DNA concentration (5 ng/µl)

favoring intramolecular ligation, followed by selection of covalently closed circular DNA (using Plasmid-Safe ATP-dependent DNase). For any particular source of genomic DNA, circularized gDNA can be prepared in advance and stored long-term at -20°C .

***In vitro* cleavage of circularized gDNA by RNP and NGS library preparation**

In vitro cleavage reaction of circularized gDNA is performed with SpCas9 ribonucleoprotein (RNPs) complexes. First, the capability of the specific RNP complex to cleave its target DNA should be tested on an amplicon containing the intended target sequence. Typical reaction conditions (1 hour incubation and relatively high 10:1 RNP:DNA molar ratio) are chosen to ensure near complete cleavage as we have previously observed that lowering the ratio of RNP:DNA reduces both fractions of amplicon cleavage and the number of CIRCLE-seq sites detected⁹. Next, RNP digested products are ligated to adapters for sequencing.

Size selection

After adapter ligation, DNA size selection (250 to 850 bp) can be optionally performed, followed by PCR. While not strictly required, DNA size selection avoids the presence of adapter dimers during PCR that might compete with the adapter ligated DNA fragments for library enrichment. Elimination of adapter dimers before sequencing is recommended to increase library diversity and sequencing quality, and to prevent loss of time and material in the case of failure or poor results due to adapter dimers.

Sample barcodes

Dual-index barcodes (NEBNext Multiplex Oligos for Illumina), are used to distinguish individual CIRCLE-seq libraries and are added by PCR in Step 74, allowing library multiplexing and increasing sample identification specificity. Ensure that combinations of dual-index barcodes are unique within sequencing runs.

Quantification and sequencing

Completed libraries are preferably quantified by droplet digital PCR²⁵ (a method that enables estimation of unwanted adapter-dimer frequencies) and sequenced with 150 bp paired-ends reads on an Illumina MiSeq instrument. Library quantification can also be performed using alternate methods such as standard quantitative PCR (qPCR)²⁵. Each pair of reads generated by Cas9 cleavage contains complete sequence information for a single on- or off-target site. CIRCLE-seq requires a modest 3-5 million sequencing reads per sample, making the technique accessible to most labs.

CIRCLE-seq data Analysis

Our freely available, open-source CIRCLE-seq data analysis pipeline (<https://github.com/tsailabSJ/circleseq>, DOI: 10.5281/zenodo.1284440) can be installed on most Mac or Linux-based operating systems. The pipeline takes as input FASTQ files and a user-generated manifest file in YAML format and produces output tables and visualization of sites of genome-wide activity of CRISPR-Cas nucleases *in vitro*.

MATERIALS

REAGENTS

sgRNA preparation

- sgRNA *in vitro* transcription plasmid pCRL1 or equivalent (Addgene)
- Oligonucleotides for sgRNA construction (IDT, see Table 2)
- SeaPlaque GTG agarose (Lonza, cat.no. 50110)
- SYBR Safe DNA gel stain (Thermo, cat.no. S33102)
- 50× TAE electrophoresis buffer (Thermo, cat.no. B49)
- Buffer QX1 (Qiagen, cat.no. 20912)
- *Bsa*I-HF (New England BioLabs, cat.no. R3535L)
- 10× CutSmart Buffer (New England BioLabs) supplied with *Bsa*I-HF
- QIAquick Gel extraction kit (Qiagen, cat.no. 28704)
- XL1-Blue subcloning grade competent cells (Agilent, cat.no. 200130)
- T4 DNA Ligase (New England BioLabs, cat.no. M0202L)
- 10× T4 DNA ligase Buffer (New England BioLabs), supplied with T4 DNA ligase
- LB agar (Miller powder) (Fisher Scientific, cat.no. DF0445174)
- Luria Broth base (Miller's LB Broth Base) (Thermo Fisher Scientific, cat.no. 244520)
- SOB broth (Fisher, cat.no. BP9737)
- Glucose (Sigma-Aldrich, cat.no. 49159-1KG)
- Kanamycin disulfate salt from *Streptomyces kanamyceticus* (Sigma, cat.no. K1876)
- QIAprep spin miniprep kit (Qiagen, cat.no. 27106)
- *Hind*III-HF (New England BioLabs, cat.no. R3104L)
- 10× CutSmart Buffer (New England BioLabs) supplied with *Hind*III-HF
- MinElute PCR purification kit (50) (Qiagen, cat.no. 28004)
- Elution buffer (Qiagen) supplied with MinElute PCR purification kit
- MEGAshortscript T7 Kit (25 rxns) (Thermo Fisher Scientific, cat.no. AM1354M)
- MEGAclear Kit 20 Rxns (Thermo Fisher Scientific, cat.no. AM1908)
- QIAxcel RNA QC Kit v2.0 (Qiagen, cat.no. 929104)
- QX RNA Size Marker 200-6000 nt (Qiagen, cat.no. 929580)

- QX RNA Alignment Marker (Qiagen, cat.no. 929510)
- RNase ZAP (Thermo Fisher Scientific, cat.no. AM9780)

Mammalian cell culture

- U2OS cells (ATCC, cat.no. HTB-96) ! **CAUTION** Check cell lines regularly for mycoplasma contamination and authenticity
- GM12878 cells (Coriell Institute, cat.no. GM12878) ! **CAUTION** Check cell lines regularly for mycoplasma contamination and authenticity
- Advanced DMEM (Thermo Fisher Scientific, cat.no. 12491023)
- RPMI 1640 (Thermo Fisher Scientific, cat.no. 72400120)
- FBS (Thermo Fisher Scientific, cat.no. 10082147)
- Penicillin-streptomycin (Thermo Fisher Scientific, cat.no. 15070063)
- Trypsin-EDTA (Thermo Fisher Scientific, cat.no. 25300120)
- PBS (Thermo Fisher Scientific, cat.no. 10010049)
- Trypan-blue (Thermo Fisher Scientific, cat.no. 15250061)

Genomic DNA isolation

- Gentra Puregene Tissue Kit (Qiagen, cat.no. 158667)
- Cell Lysis Solution (Qiagen) supplied with Gentra Puregene Kit
- Protein Precipitation Solution (Qiagen) supplied with Gentra Puregene Kit
- DNA Hydration Solution (Qiagen) supplied with Gentra Puregene Kit
- Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, cat.no. Q32853)
- Isopropanol (Sigma-Aldrich, cat.no. 190764) ! **CAUTION** Isopropanol is flammable, keep away from open flame.
- RNase A (Qiagen, cat.no. 19101)
- Proteinase K (Qiagen, cat.no. 19131)

Genomic DNA shearing

- IDTE pH 8.0 (1× TE Solution) (Integrated DNA Technologies, cat.no. 11050204)
- Agencourt AMPure XP - 60 ml (Beckman Coulter, cat.no. A63881)

CIRCLE-seq library preparation and NGS

- Phusion Hot Start Flex 2× Master Mix (New England BioLabs, cat.no. M0536L)
- HTP Library Preparation Kit PCR-free (96rxn) (Kapa Biosystems, cat.no. KK8235)

- PEG/NaCl SPRI solution, supplied with HTP Library Preparation Kit PCR-free (96rxn) (Kapa Biosystems)
- 2× Kapa HiFi HotStart Ready Mix (Kapa Biosystems, cat.no. KK2602)
- T4 Polynucleotide Kinase (PNK) (New England BioLabs, cat.no. M0201L)
- T4 DNA Ligase (New England BioLabs, cat.no. M0202L)
- 10× T4 DNA ligase Buffer (New England BioLabs), supplied with T4 DNA Ligase
- USER Enzyme (New England BioLabs, cat.no. M5505L)
- Exonuclease I (*E. coli*) (New England BioLabs, cat.no. M0293L)
- Lambda Exonuclease (New England BioLabs, cat.no. M0262L)
- Plasmid-Safe ATP-dependent DNase (Epicentre, cat.no. E3110K)
- Plasmid-Safe 10× Reaction Buffer (Epicentre), supplied with Plasmid-Safe ATP-dependent DNase
- 25mM ATP solution (Epicentre), supplied with Plasmid-Safe ATP-dependent DNase
- Cas9 nuclease *S. pyogenes* (New England BioLabs, cat.no M0386M)
- 10× Cas9 buffer (New England BioLabs), supplied with Cas9 nuclease *S. pyogenes*
- NEBNext® Multiplex Oligos for Illumina® (Dual Index Primers Set 1) (New England BioLabs, cat.no. E7600S)
- NEBNext adapter for Illumina (New England BioLabs), supplied with NEBNext® Multiplex Oligos for Illumina®
- Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, cat.no. Q32853)
- Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, cat.no. Q32854)
- Qubit assay tubes (Thermo Fisher Scientific, cat.no. Q32856)
- ddPCR™ SuperMix for probes (Bio-Rad, cat.no. 1863010)
- Droplet Generation Oil for Probes (Bio-Rad, cat.no. 1863005)
- Droplet Reader Oil (Bio-Rad, cat.no. 1863004)
- IDTE pH 8.0 (1× TE Solution) (Integrated DNA Technologies, cat.no. 11050204)
- MiSeq® Reagent Kit v3 (600 cycle) (Illumina, cat.no. MS-102-3003)
- Flow Cell, supplied with MiSeq® Reagent Kit v3 (600 cycle)
- Hyb Buffer, supplied with MiSeq® Reagent Kit v3 (600 cycle)
- PhiX Control V3 KIT (Illumina, cat.no. FC-110-3001)

- Sodium hydroxide solution, volumetric, 1 M NaOH (1N) (Sigma-Aldrich, cat.no. 71463-1L) ! **CAUTION** Sodium hydroxide is hazardous in case of contact with skin, contact with eyes, inhalation and ingestion. Wear properly protective clothing and handle with care.
- North Alcohol Wipes (Dynarex, cat.no. 19-014-855)
- VWR Lens Cleaning Tissue (VWR, cat.no. 52846-001)
- Agencourt AMPure XP - 60 ml (Beckman Coulter, cat.no. A63881)
- Ethanol (Sigma, cat.no. E7023) ! **CAUTION** Ethanol is flammable, keep away from open flame.
- QIAxcel DNA High Resolution Kit (Qiagen, cat.no. 929002)
- QX Size Marker 100 bp – 2.5 kb (Qiagen, cat.no. 929559)
- QX Alignment Marker 15 bp – 5 kb (Qiagen, cat.no. 929524)
- Tween-20 (Sigma-Aldrich, cat.no. P7949)
- Tris Base (Fisher, cat.no. BP1521)
- HCl (Fisher, cat.no. A144500) ! **CAUTION** Hydrochloric acid is hazardous in case of contact with skin, contact with eyes, inhalation and ingestion. Wear properly protective clothing and handle with care.
- EDTA 0.5 M (Thermo Fisher Scientific, cat.no. 15575020)
- 1.5% PippinHT Cassette (Sage Science, cat.no. HTC1510)
- Primers, probes and adapters - sequences are listed in Table 2 (Integrated DNA Technologies)

EQUIPMENT

- Covaris E220 (Covaris)
- microTUBE AFA Fiber Pre-Slit Snap-Cap 6×16mm (25) (Covaris, cat.no. 520045)
- DG8 Gaskets (Bio-Rad, cat.no. 1863009)
- DG8 Cartridges (Bio-Rad, cat.no.1864008)
- DG8 Cartridge holder (Bio-Rad, cat.no. 1863051)
- Twin.tec PCR Plate 96, semi-skirted, green (Eppendorf, cat.no. E951020346)
- Pierceable Foil Heat Seal (Bio-Rad, cat.no. 1814040)
- QIAxcel (Qiagen, cat.no. 9001941)
- Qubit fluorometer (Thermo Fisher Scientific, cat.no. Q33226)
- Qubit assay tubes (Thermo Fisher Scientific, cat.no. Q32856)

- Thermocycler with programmable temperature stepping functionality, 96 well (Applied Biosystems Veriti, cat.no. 4375786)
- QX200™ Droplet Digital™ PCR System (Bio-Rad, cat.no. 1864001, containing a QX200™ Droplet Generator and a QX200™ Droplet Reader)
- MiSeq System (Illumina, cat.no. SY-410-1003)
- Magnum FLX Enhanced Universal Magnet Plate (Alpaqua, cat.no. A00400)
- Tissue culture dish 150×25mm (Fisher, cat.no. 0877224)
- T-75 flasks (Fisher, cat.no. 7202000)
- Filter unit 500 ml (Fisher, cat.no. SCGPU05RE)
- Petri dishes (Fisher, cat.no. FB0875713)
- Hemocytometer disposable (Thermo Fisher Scientific, cat.no. 22600100)
- 0.2 ml thin-walled 12 tube and flat cap strips (Thermo Fisher Scientific, cat.no. AB1114)
- NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific, cat.no. ND-8000-GL)
- PippinHT (Sage Science)

Software

- Python 2.7: <https://www.anaconda.com/download/>
- BWA: <https://github.com/lh3/bwa>
- Samtools: <http://www.htslib.org/download/>
- CIRCLE-seq: <https://github.com/tsailabSJ/circleseq> (DOI: 10.5281/zenodo.1284440)

General laboratory consumables

- Filtered sterile pipette tips
- Microcentrifuge tubes, 1.5 ml (Axygen, cat.no. 31104051)
- Semi-skirted PCR plates (Thermo Fisher Scientific, 14230244)
- Conical tubes 15 ml (Fisher, cat.no. 339651)
- Conical tubes 50 ml (Fisher, cat.no. 339653)
- Serological pipettes 5 ml (Fisher, cat.no. 170355)
- Serological pipettes 10 ml (Fisher, cat.no. 12567603)
- Serological pipettes 25 ml (Fisher, cat.no. 12567604)
- Reagent reservoir 25 ml (Fisher cat.no.2138127C)

General laboratory equipment

- Bacterial incubator: Precision low temperature BOD refrigerated incubator (Thermo Fisher Scientific) or equivalent
- Bacterial shaker: Multitron Pro Shaker (Infors-HT) or equivalent
- Cell culture incubator: Heracell VIOS 160i CO₂ incubator (Thermo Fisher Scientific) or equivalent
- Benchtop microcentrifuge (Eppendorf, cat.no. 05400002)
- Eppendorf ThermoMixer (Eppendorf, cat.no. 5384000020)
- Inverted Microscope

REAGENT SETUP

10× STE annealing buffer. Prepare buffer as follows:

Component	Amount (ml)	10× concentration
Tris-HCl, pH 8.0, 1M	5	100 mM
NaCl 5M	5	500 mM
EDTA 0.5M	1	10 mM
H ₂ O	39	
Total	50	

10× STE annealing buffer can be stored at room temperature (25 °C) for several months.

Ethanol, 70% (vol/vol) Prepare 70% (vol/vol) ethanol in UltraPure water right before use.

Ethanol, 80% (vol/vol) Prepare 80% (vol/vol) ethanol in UltraPure water right before use.

NaOH, 0.2 N Dilute 20 µl of NaOH 1 M in 80 µl of UltraPure water. Make it fresh right before use.

Tris-HCl 10 mM + 0.1% Tween-20 Prepare a solution with 10 mM Tris-HCl (pH 8.0), 0.1% Tween-20 (vol/vol), as follows:

Component	Amount (ml)	Final concentration
Tris-HCl, pH 8.0, 1M	0.1	10 mM
Tween-20	0.01	0.1% (vol/vol)
H ₂ O	9.89	
Total	10	

Tris-HCl 10 mM + 0.1% Tween-20 can be stored at room temperature for several months.

U2OS culture medium For culture of human U2OS cells, prepare Advanced DMEM medium supplemented with 1× GlutaMAX, 10% FBS (vol/vol) and 1× penicillin-streptomycin, and sterile filter using a 500 ml filter unit. Store the medium at 4 °C for up to 1 month.

GM12878 culture medium For culture of human GM12878 cells, prepare RPMI 1640 medium supplemented with 10% FBS (vol/vol) and 1× penicillin-streptomycin, and sterile filter using a 500 ml filter unit. Store the medium at 4 °C for up to 1 month.

0.1× elution buffer Dilute elution buffer in UltraPure water to a 0.1× working solution, and store at room temperature for several months.

SOC medium Reconstitute one capsule of SOB in 500 ml of water. Autoclave the mixture to sterilize. Allow the SOB to cool down before and add glucose to a final concentration of 20mM. SOC medium can be stored at room temperature for up to 3 months.

LB broth Reconstitute LB broth (Luria Broth base) at a concentration of 25 g liter⁻¹ in deionized water and swirl to mix. Autoclave the mixture to sterilize. Allow the LB to cool down before adding kanamycin to a final concentration of 50 µg/ml. LB broth can be stored at room temperature for up to 3 months.

LB agar plates (100-mm petri dish, kanamycin) Reconstitute the LB agar at a concentration of 32 g liter⁻¹ in deionized water and swirl to mix. Autoclave the mixture to sterilize. Allow the LB agar to cool down to approximately 65 °C before adding kanamycin to a final concentration of 50 µg/ml. Pour 25 ml of LB agar per 100-mm dish. Place the lids on the plates and allow them to cool down until solidified. Invert the plates and let them sit overnight. Agar plates can be stored in plastic bags at 4 °C for up to 3 months.

1× TAE electrophoresis buffer Dilute 50× TAE electrophoresis buffer to 1× in distilled water, using 20 ml of 50× TAE electrophoresis buffer and 980 ml of distilled water. Typically, 1× TAE is made up right before use but can be stored indefinitely.

PROCEDURE

CRITICAL: All reactions (from Steps 9-84) are performed in PCR plates and when in the thermocycler, they are incubated with heated lid.

Cell culture ● TIMING 5 d

1| Culture U2OS cells using Option A; GM12878 cells using Option B; or any cells of interest using Option C.

(A) U2OS maintenance culture.

- (i) Culture cells in Advanced DMEM medium supplemented with 10% (vol/vol) FBS, 1× GlutaMAX, and 1× penicillin-streptomycin at 37 °C and 5% CO₂ in 150 mm tissue culture plates.
- (ii) To passage, remove the medium and rinse once with 25 ml of PBS. Add 5 ml of pre-warmed Trypsin-EDTA and incubate at 37 °C for 5

min. Add 20 ml of prewarmed medium to inactivate trypsin and transfer the cells to a 50 ml conical tube. Centrifuge at room temperature for 5 min at $300 \times g$ and discard the supernatant. Re-suspend the cells in fresh pre-warmed Advanced DMEM medium (prepared as described in Reagent Setup U2OS cell culture medium) by pipetting them up and down gently, and then, reseed them into new plates.

▲ **CRITICAL STEP** Avoid letting cells become fully confluent by passaging every 3-4 days at a split ratio of 1:3 to 1:8.

- (iii) Harvest the cells at 90% confluence (as observed under a standard inverted microscope). Wash the plate with 25 ml of PBS, add 5 ml of warmed Trypsin-EDTA to a 150-mm plate, and incubate at 37 °C for 5 min. Add 20 ml of warmed complete medium to inactivate trypsin, and transfer the cells to a 50 ml conical tube. Centrifuge at room temperature for 5 min at $300 \times g$ and discard the supernatant. Re-suspend the cells in 10 ml of PBS. Take a 10 μ l aliquot of the cells resuspended in PBS, mix 1:1 with trypan-blue and count in a hemocytometer. Aliquot 2×10^7 cells/tube, centrifuge at room temperature for 3 min at $5000 \times g$, and discard supernatant.

■ **PAUSE POINT** Cell pellet(s) can be frozen at -20 °C for several months.

(B) GM12878 maintenance culture.

- (i) Culture cells in RPMI-1640 medium supplemented with 10% (vol/vol) FBS and $1 \times$ penicillin-streptomycin at 37 °C and 5% CO₂ in T-75 flasks.
- (ii) As these cells grow in suspension, to passage, collect the medium containing the cells into 50 ml conical tubes, centrifuge at room temperature for 10 min at $300 \times g$ and discard the supernatant. Re-suspend the cells in fresh pre-warmed RPMI-1640 medium (prepared as described in Reagent Setup GM12878 cell culture medium) by pipetting them up and down gently, and then reseed them into new T75 flasks, splitting in 1:2 to 1:3 ratio.
- (iii) Harvest the cells by collecting the medium containing the cells into 50 ml conical tubes, centrifuge at room temperature for 10 min at $300 \times g$ and discard the supernatant. Re-suspend the cells in 10 ml of PBS and count in a hemocytometer. Aliquot 2×10^7 cells/tube, spin down at $5000 \times g$ for 3 min. Aspirate PBS.

■ **PAUSE POINT** Cell pellet can be frozen at -20 °C for several months.

(C) Culture cells of interest.

- (i) Culture cells of interest according to pre-established protocols, harvest and resuspend in 10 ml of PBS, take a 10 μ l aliquot of the cells re-suspended in PBS, mix 1:1 with trypan-blue and count in a hemocytometer. Aliquot 2×10^7 cells/tube, spin down at $5000 \times g$ for 3 min.

■ **PAUSE POINT** Cell pellet can be frozen at -20°C for several months.

Genomic DNA isolation ● TIMING 1 d

2| Perform genomic DNA isolation with Gentra Puregene Kit (Qiagen), following the manufacturer's instructions.

3| Re-suspend the cell pellet in 200 μ l of PBS in a 15 ml conical tube, add 3 ml of Cell Lysis Solution (supplied with Gentra Puregene Kit) and 15 μ l of Proteinase K, mix by inverting 25 times, and incubate at 55°C with shaking for 3 h or overnight for maximum yield.

4| Add 15 μ l of RNase A solution, mix by inverting 25 times. Incubate at 37°C for 1 h.

5| Incubate at 1 min on ice and add 1 ml of Protein Precipitation Solution (provided with Gentra Puregene Kit (Qiagen)). Vortex vigorously for 20 s at high speed. Centrifuge at room temperature for 10 min at $2000 \times g$. The precipitated proteins should form a tight pellet. If the protein pellet is not tight, incubate on ice for 5 min and repeat the centrifugation.

6| Pipet 3 ml of isopropanol into a clean 15 ml conical tube and add the supernatant from the previous step by pipetting carefully. Mix by inverting 50 times. Centrifuge at room temperature for 3 min at $2000 \times g$. Carefully discard the supernatant and drain the tube by inverting on a clean piece of absorbent paper, taking care that the DNA pellet remains in the tube.

▲ **CRITICAL STEP** Make sure the protein pellet is not dislodged from the original tube during removal of the supernatant, and discard original tube with protein pellet once the supernatant has been collected.

7| Add 3 ml of 70% ethanol and invert several times to wash the DNA pellet and centrifuge at room temperature for 1 min at $2000 \times g$. Discard the supernatant.

8| Air dry for 5-10 min. Add 400 μ l of DNA Hydration Solution (supplied with Gentra Puregene Kit (Qiagen)), mix well by pipetting up and down. Incubate at 65°C for 1 h to dissolve the DNA. Incubate at room temperature overnight. Briefly centrifuge at room temperature for 1 min at $2000 \times g$ and quantify genomic DNA by Qubit fluorometer using Qubit dsDNA BR assay and proper Qubit assay tubes.

▲ **CRITICAL STEP** Do not over-dry the pellet as the DNA will be hard to dissolve.

■ **PAUSE POINT** Isolated genomic DNA can be stored at -20°C for several months.

Preparation of sgRNA ● TIMING 1 w

9| To obtain sgRNA, either have the synthetic sgRNA commercially synthesized (option A) or perform *in vitro* transcription from a plasmid template such as pCRL1 or equivalent (option B).

▲ **CRITICAL STEP** It is important to use the same sgRNA in subsequent cell-based genome editing experiments carried out to confirm *in vitro* CIRCLE-seq results.

(A) Commercially synthesized sgRNA.

- (i) Order synthetic sgRNA from various commercial sources (i.e. Synthego, Trilink, IDT).

(B) *In vitro* transcribed sgRNA.

- (i) Re-suspend each oligonucleotide for each sgRNA to a final concentration of 100 μM in 1 $\times$ TE and prepare the following mix to anneal oligonucleotides:

Component	Volume (μl)	Final concentration
sgRNA Top 100 μM	10	10 μM
sgRNA Bottom 100 μM	10	10 μM
STE 10 $\times$	10	1 $\times$
Nuclease-free H_2O	70	
Total	100	

- (ii) Anneal using the following program in a thermocycler: 95 $^{\circ}\text{C}$ for 5 min, $-1^{\circ}\text{C}/\text{min}$ for 70 cycles, hold at 4 $^{\circ}\text{C}$.

■ **PAUSE POINT** annealed oligonucleotides can be stored at -20°C for several months.

- (iii) Dilute the annealed oligo 1:1000 by adding 1 μl of oligo to 999 μl of TE.
- (iv) To generate the sgRNA expression construct, clone annealed oligos into a sgRNA expression plasmid such as pCRL1. Digest 3-5 μg of pCRL1 backbone plasmid with *BsaI*-HF, as follows:

Component	Volume (μl)	Final concentration
CutSmart buffer (10 $\times$)	5	1 $\times$
pCRL1 vector backbone (3-5 μg)	variable	variable
SsaI-HF (20 U/ μl)	2	0.8 U/ μl
Nuclease-free H_2O	to 50	
Total	50	

Incubate in a thermocycler at 37 °C for 2 h, hold at 4 °C.

- (v) Run the entire volume of the SsaI-digested backbone in a 1% low melt agarose gel (SeaPlaque GTG agarose with SYBR Safe dye) in 1× TAE buffer at 5 V/cm for 2 h. Cut the digested backbone from the gel, transfer to a 1.5 ml tube and incubate at 65 °C for 10 min or until the gel is completely melted. Add 3 volumes of Qiagen Buffer QX1 and incubate at 50 °C for 10 min. Add 1 volume of isopropanol and transfer the mixture to a QIAquick gel purification column (supplied with QIAquick gel extraction kit) and centrifuge at room temperature for 1 min at $17,900 \times g$. Discard the liquid and add 750 µl of 70% ethanol. Centrifuge at room temperature for 1 min at $17,900 \times g$. Repeat for a total of two ethanol washes. Perform an additional centrifugation step at room temperature for 3 min at $17,900 \times g$ to eliminate excess ethanol. Elute the DNA in 20 µl of 0.1× elution buffer by centrifuging at room temperature for 1 min at $17,900 \times g$. Quantify by NanoDrop.
- (vi) To clone the sgRNA oligoduplex into pCRL1 (or equivalent) *in vitro* transcription vector, set up a standard ligation reaction for each sgRNA, as follows:

Component	Volume (µl)	Final concentration
T4 DNA ligase buffer (10×)	1	1×
BsaI-digested vector backbone 20 ng	variable	2 ng/µl
Diluted oligoduplex from (iii)	2	2 nM
T4 PNK (10 U/µl)	0.5	0.5 U/µl
T4 DNA ligase (400 U/µl)	0.5	20 U/µl
Nuclease-free H ₂ O	to 10	
Total	10	

▲ CRITICAL STEP: We recommend also preparing a no-insert vector-only negative control for ligation, to estimate the likely ratio of correct to incorrect clones.

- (vii) Incubate the ligation reaction at 16 °C for 30 min.
- (viii) Transform the ligation into XL1blue competent *E. coli* strain, according to the protocol supplied with the cells. Briefly, add 5 µl of the ligation into 50 µl of ice-cold chemically competent XL1-Blue cells, incubate the mixture on ice for 15-30 min, heat-shock at 42 °C for 45 s and return to ice for 2 min. Add 450 µl of SOC medium and incubate at 37 °C with shaking at 250 rpm for 1 h. Plate onto a LB-agar plate containing 50 µg/ml kanamycin. Incubate it overnight at 37 °C.

- (ix) Next day: check the plates for colony growth. Usually, there are few to no colonies on the negative control plates and many colonies on the sgRNA cloned plates.

? TROUBLESHOOTING

- (x) Pick two to four colonies to check the insertion of sgRNA target. Use a sterile pipette tip to inoculate a single colony into 4-5 ml of LB medium with 50 µg/ml kanamycin. Incubate the culture at 37 °C overnight with shaking at 250 rpm.
- (xi) Next day: isolate plasmid DNA using QIAprep spin miniprep kit (Qiagen) according to the manufacturer's instruction.

■ **PAUSE POINT** Isolated plasmid can be stored at –20 °C for several months.

- (xii) Sequence validation of correct clones. Check the sequence of each plasmid by sequencing using M13 reverse primer. Align the sequencing results against the pCRL1 cloning vector sequence to verify the 20-nt sequence inserted between the T7 promoter and the sgRNA scaffold is correct.

? TROUBLESHOOTING

- (xiii) Setup a reaction to digest 3-5 µg of sequence-verified sgRNA plasmid with *Hind*III-HF for subsequent *in vitro* transcription as follows:

Component	Volume (µl)	Final concentration
CutSmart buffer (10×)	5	1×
sgRNA plasmid (3-5 µg)	Variable	variable
<i>Hind</i> III-HF (20 U/µl)	4	1.6 U/µl
Nuclease-free H ₂ O	to 50	
Total	50	

- (xiv) Incubate in a thermocycler at 37 °C for 2 h. Heat inactivate the restriction enzyme at 85 °C for 20 min, hold at 4 °C.
- (xv) Purify the linearized plasmid using MinElute PCR purification kit, following the manufacturer's instructions and elute in 12 µl of 0.1× elution buffer.

▲ **CRITICAL STEP** Use RNase Free water and clean bench and pipettes with RNase Zap before starting and keep everything RNase-free.

- (xvi) Set up the *in vitro* transcription reaction using the MEGAshortscript T7 kit as follows:

Component	Volume (μl)	Final concentration
T7 10× Reaction Buffer	2	1×
<i>Hind</i> III Digested DNA	8	variable
dATP (75 mM)	2	7.5 mM
dUTP (75 mM)	2	7.5 mM
dCTP (75 mM)	2	7.5 mM
dGTP (75 mM)	2	7.5 mM
Enzyme Mix	2	
Total	20	

- (xvii) Incubate the *in vitro* transcription reaction in a thermocycler at 37 °C for 16 h. Add 1 μl of DNase I (2 U/μl) and incubate at 37 °C for 15 min.
- (xviii) Clean up the *in vitro* transcribed product with the MEGAclean purification kit following the manufacturer's instruction. Elute sgRNA with 50 μl of prewarmed (95 °C) elution buffer. After the centrifugation step, collect the eluted product and re-load into the column in order to perform a second elution step to increase the yield. Quantify the sgRNA by NanoDrop. Run 100 ng of each sgRNA on QIAxcel instrument for sgRNA quality control, using QIAxcel RNA QC Kit v2.0 (Qiagen), QX RNA Size Marker 200-6000 nt (Qiagen) and QX RNA Alignment Marker (Qiagen), following manufacturer's instructions.
- **PAUSE POINT** Make 5 μl aliquots and freeze at –80 °C. sgRNA aliquots can be stored at –80 °C for several months.

sgRNA *In vitro* cleavage test ● TIMING 5 h

CRITICAL: Note that in this protocol, we use a target in the *EMX1* gene as an example target *locus*. For other targets, appropriate primers to amplify the target of interest should be used in place of the example *EMX1* primers.

10| Assemble PCR reaction as follows:

Component	Volume (μl)	Final concentration
Phusion Hot Start Flex 2× Master Mix	25	1×
oSQT1143 F 10 μM (Table 2)	0.5	0.1 μM
oSQT1144 R 10 μM (Table 2)	0.5	0.1 μM
gDNA (U2OS gDNA, 100 ng, 20 ng/μl, from Step 8)	5	2 ng/μl
H ₂ O	19	
Total	50	

11| Perform the PCR using the following thermocycling conditions to amplify on-target *locus*. Optimize conditions as required to obtain a single, sharp band when analyzed on a QIAxcel instrument or standard gel electrophoresis.

Step	Temperature	Time	Cycles
Denaturation	98 °C	2 min	1
Denaturation	98 °C	10 s	10
Annealing	72-62 °C (–1 °C/cycle)	15	10
Extension	72 °C	30 s	10
Denaturation	98 °C	10 s	30
Annealing	62 °C	15 s	30
Extension	72 °C	30 s	30
Hold	4 °C		1

12| Purify the PCR product with Agencourt AMPure XP beads. Perform cleanup according to manufacturer's instructions. Briefly, add 1.8× volumes (90 µl) of Agencourt AM-Pure XP beads to the PCR, mix thoroughly by pipetting 10 times. Incubate at room temperature for 5 min. Place the reaction plate onto a Magnum FLX magnetic rack for 3 min. Remove the cleared solution from the reaction plate and discard. Add 200 µl of 80% ethanol, incubate for 30 s and remove the supernatant. Repeat this step for a total of two ethanol washes. Remove ethanol completely and let the samples air dry for 3 min on the magnetic rack. Remove the plate from the magnetic rack and add 40 µl of TE pH 8.0, and pipette 10 times to mix. Incubate at room temperature for 2 min. Place the reaction plate back to the magnetic rack for 1 min. Transfer the supernatant to a new plate. Quantify the purified PCR by NanoDrop and run on a QIAxcel capillary electrophoresis instrument, in a 0.2 ml thin-walled 12-well strip tube with a QIAxcel DNA High Resolution Kit (Qiagen), QX Alignment Marker 15 bp – 5 kb (Qiagen) and QX Size Marker 100 bp – 2.5 kb (Qiagen), following manufacturer's instruction.

■ **PAUSE POINT** Purified PCR product can be stored at –20 °C for several months.

13| Dilute Cas9 nuclease to 1 µM as follows:

Component	Volume (µl)	Final concentration
Cas9 Nuclease Reaction Buffer (10×)	2	1×
Cas9 Nuclease, <i>S. pyogenes</i> (20 µM)	1	1 µM
H ₂ O	17	
Total volume	20	

14| Dilute sgRNA (from Step 9) to 3 µM in H₂O to a total volume of 10µl. The following formula can be used to calculate the approximate molecular weight of sgRNA:

$$\text{molecular weight ssRNA (g / mol)} = (\text{length of ssRNA (nt)} \times 321.47 \text{ g / mol}) + 18.02 \text{ g / mol}$$

For a typical sgRNA of length 104 nt, 3 μM is approximately 100 ng/ μL .

▲ **CRITICAL STEP** Use RNase-free technique to avoid degrading sgRNA.

15| Set up the *in vitro* cleavage reaction as follows:

Component	Volume (μL)	Final concentration
Cas9 Nuclease Reaction Buffer (10 $\times$)	5	1 $\times$
Cas9 Nuclease, <i>S. pyogenes</i> (1 μM)	4.5	90 nM
<i>In vitro</i> transcribed sgRNA (3 μM)	1.5	90 nM
Total cleavage master-mix	11	

16| Incubate at room temperature for 10 min to complex sgRNA with Cas9 protein.

17| Add the purified PCR product (from Steps 10-12), diluted in H_2O to a volume of 39 μL . The final reaction volume is 50 μL .

Component	Volume (μL)	Final concentration
Cleavage master mix	11	
Purified PCR product (125 ng)	39	2.5 ng/ μL
Total	50	

18| Incubate in a thermocycler at 37 $^{\circ}\text{C}$ for 1 h, hold at 4 $^{\circ}\text{C}$.

19| Add 1 μL of Proteinase K to the *in vitro* cleaved DNA and incubate at 37 $^{\circ}\text{C}$ for 15 min.

20| Purify the *in vitro* cleaved, proteinase K-treated DNA with 1 $\times$ (50 μL) volumes of Agencourt AMPure XP beads according to manufacturer's instruction (summarized in Step 12) and elute in 20 μL of 0.1 $\times$ EB.

21| Run the purified *in vitro* cleaved DNA on a QIAxcel capillary electrophoresis instrument, side-by-side with the uncleaved amplicon (purified PCR product from Steps 10-12). Proceed to next steps after confirming that PCR product is cleaved to near completion (80%).

? TROUBLESHOOTING

DNA shearing ● TIMING 3h

22| Covaris E220 shearing instrument: Fill the tank with purified water up to “run” level 5. Open the software and click “ON” to degas and setup the temperature to 4.5 °C. Wait 1 hour for the degassing process to be completed before starting shearing.

23| Aliquot genomic DNA into Covaris microTUBE AFA Fiber Pre-Slit Snap-Cap 6×16mm tubes. Each tube should contain 25 µg of genomic DNA and be filled up to a total volume of 130 µl with 1× TE.

▲ CRITICAL STEP Volume must be 130 µl to avoid creation of bubbles that interfere with acoustic shearing.

24| Shear the DNA to an average length of 300 bp, with the following parameters:

Target bp (Peak)	300
Peak incident power (W)	140
Duty Factor	10%
Cycles per Burst	200
Treatment Time (s)	80

Purification of sheared genomic DNA ● TIMING 1 h

25| Purify and quantify sheared genomic DNA. Divide the sheared genomic DNA in two aliquots (65 µl) and purify with 1.8× (117 µl) volumes of AMPure XP beads as described in Step 12. Elute the sheared gDNA with 50 µl of TE. Transfer the supernatants to a new plate. Pool and quantify by Qubit dsDNA BR assay and proper Qubit assay tubes.

26| Run 2 µl of the sheared gDNA on a QIAxcel capillary electrophoresis instrument, following manufacturer’s instruction, to check the size distribution. A broad distribution with a median size of 300 bp is expected.

■ PAUSE POINT Sheared genomic DNA without beads can be stored at –20 °C for several months.

CIRCLE-seq library preparation ● TIMING 3 d

CRITICAL: see Fig. 2 for a detailed schematic overview of CIRCLE-seq library preparation method.

27| Hairpin adapter annealing. Re-suspend the hairpin adapter oSQT1288 (Table 2) to a final concentration of 100 µM in 1× TE. Perform the adapter annealing as follows:

Component	Volume (µl)	Final concentration
oSQT1288 100 µM	40	40 µM

Component	Volume (μl)	Final concentration
STE 10×	10	1×
Nuclease-free H ₂ O	50	
Total	100	

28| Annealing program: 95 °C for 5 min, −1 °C /min for 70 cycles, hold at 4 °C.

The annealed adapter will be necessary in Step 37 and can be stored at −20 °C for several months.

29| End Repair. Setup the end-repair master-mix from the KAPA HTP Library Preparation Kit PCR-free kit:

Component	Volume (μl)	Final concentration
Nuclease-free H ₂ O	8	
Kapa End Repair Buffer (10×	7	1×
Kapa End Repair Enzyme Mix	5	
Total End-repair master-mix	20	

30| Add 20 μl of the mix to the sheared genomic DNA from Steps 24-25.

Component	Volume (μl)	Final concentration
End-repair mix	20	
Sheared genomic DNA (3-5 μg)	50	<i>variable</i>
Total	70	

31| Incubate in a thermocycler at 20 °C for 30 min, hold at 4 °C.

32| Add 1.7× volumes (120 μl) of Agencourt AMPure XP beads and purify as previously described in Step 12. Elute by adding 42 μl of TE pH 8.0, keeping the beads in solution for the following reaction.

▲ CRITICAL STEP Keep the beads in solution for the next enzymatic step.

33| A-tailing. Setup the A-tailing master-mix:

Component	Volume (μl)	Final concentration
Kapa A-tailing Buffer (10×	5	1×
Kapa A-tailing Enzyme	3	

Component	Volume (μl)	Final concentration
Total A-tailing master-mix	8	

34| Add 8 μl of A-tailing master-mix to each eluted DNA sample with beads from Step 32.

Component	Volume (μl)	Final concentration
A-tailing master-mix	8	
End-repaired DNA/beads	42	<i>variable</i>
Total	50	

35| Incubate in a thermocycler at 30 °C for 30 min, hold at 4 °C.

36| Add 1.8× volumes (90 μl) of PEG/NaCl SPRI solution (supplied with HTP Library Preparation Kit PCR-free (96rxn), Kapa Biosystems) to the A-tailed DNA and purify the A-tailed DNA as previously described in Step 12. Elute in 30 μl of TE pH 8.0, keeping the beads in solution for the following reaction.

▲ CRITICAL STEP Keep the beads in solution for the next enzymatic step.

37| Adapter ligation. Setup the Adapter-ligation master-mix:

Component	Volume (μl)	Final concentration
Kapa Ligation Buffer (5×)	10	1×
Kapa DNA Ligase	5	
Annealed Hairpin Adapter (40 μM) from Step 27	5	4 μM
Total Adapter-ligation master-mix	20	

38| Add 20 μl of Adapter-ligation master-mix to each eluted DNA sample with beads.

Component	Volume (μl)	Final concentration
Adapter-ligation master-mix	20	
A-tailed DNA/beads	30	<i>variable</i>
Total	50	

39| Incubate in a thermocycler at 20 °C for 1 h, hold at 4 °C.

40| Add 1× volumes (50 μl) of PEG/NaCl SPRI solution to the adapter-ligated DNA and purify the adapter-ligated DNA as previously described in Step 12. Elute with 30 μl of TE

pH 8.0. Transfer the supernatants to a new plate. Pool and quantify by Qubit dsDNA BR assay using proper Qubit assay tubes.

■ **PAUSE POINT** Purified adapter ligated DNA can be stored at -20°C for up to one month according to manufacturer's recommendations (Kapa HTP Library Preparation Kit PCR-free).

41| Setup the Lambda Exonuclease and Exonuclease I (*E. coli*) reaction master-mix, which is used to eliminate DNA molecules (linear single- or double-stranded DNA) that do not have the adapters ligated to both ends:

Component	Volume (μl)	Final concentration
Exonuclease I Reaction Buffer (10 $\times$)	5	1 $\times$
Lambda Exonuclease (5 U/ μl)	4	0.4 U/ μl
Exonuclease I (<i>E. coli</i>) (20 U/ μl)	1	0.4 U/ μl
Total Lambda Exo/ExoI master-mix	10	

42| Add 10 μl of Lambda Exonuclease and Exonuclease I master-mix to 1 μg of adapter-ligated dNa.

Component	Volume (μl)	Final concentration
Lambda Exo and Exo I master-mix	10	
Adapter ligated DNA (1 μg)	40	20 ng/ μl
Total	50	

43| Incubate in a thermocycler at 37°C for 1 h, 75°C for 10 min, hold at 4°C .

44| Add 1.8 $\times$ volumes (90 μl) of AMPure XP beads to the Lambda exonuclease and Exonuclease I treated DNA. Purify as described in Step 12. Elute in 40 μl of TE pH 8.0, keeping the beads in solution.

▲ **CRITICAL STEP** Keep the beads in solution for the next enzymatic step.

45| USER enzyme and T4 Polynucleotide Kinase (PNK) treatment. This step is required for releasing 4 bp overhangs and preparing ligation-compatible DNA ends required for ligation in the subsequent reaction (Step 49). Setup USER enzyme and T4 Polynucleotide Kinase master-mix:

Component	Volume (μl)	Final concentration
T4 DNA Ligase Buffer (10 $\times$)	5	1 $\times$

Component	Volume (μl)	Final concentration
USER Enzyme (1 U/μl)	3	0.06 U/μl
T4 PNK (10 U/μl)	2	0.4U/μl
Total USER and PNK master-mix	10	

46| Add 10 μl of USER enzyme and T4 PNK master-mix to each Lambda and Exonuclease I treated DNA sample with beads.

Component	Volume (μl)	Final concentration
USER and PNK master-mix	10	
Lambda exonuclease and Exonuclease I treated DNA and beads	40	<i>variable</i>
Total	50	

47| Incubate in a thermocycler at 37 °C for 1 h, hold at 4 °C.

48| Add 1.8× volumes (90 μl) of PEG and NaCl SPRI solution to the USER and T4 PNK treated DNA and purify as described in Step 12. Elute in 35 μl of TE pH 8.0. Transfer the supernatant to a new plate. Pool and quantify by Qubit dsDNA HS assay using proper Qubit assay tubes.

49| Intramolecular circularization. Setup circularization master-mix:

Component	Volume (μl)	Final concentration
Nuclease-free water	8	
T4 DNA Ligase Buffer (10×)	10	1×
T4 DNA Ligase (400 U/μl)	2	8 U/μl
Total Circularization master-mix	20	

50| Add 20 μl of circularization master-mix to 500 ng of USER/PNK treated DNA.

Component	Volume (μl)	Final concentration
Circularization master-mix	20	
USER/PNK treated DNA (500 ng)	80	5 ng/μl
Total	100	

51| Incubate in a thermocycler at 16 °C for 16 h (overnight).

52| Add 1× volumes (100 µl) of AMPure XP beads to the circularized DNA and purify as described in Step 12. Elute in 38 µl of TE pH 8.0. Transfer the supernatant to a new plate.

53| Exonuclease degradation of residual linear DNA (Plasmid-Safe ATP-dependent DNase Treatment). Setup Plasmid-Safe ATP-Dependent DNase Treatment master-mix:

Component	Volume (µl)	Final concentration
Plasmid-Safe Reaction Buffer (10×)	5	1×
ATP (25 mM)	2	1 mM
Plasmid-Safe ATP-Dependent DNase (10 U/µl)	5	1 U/µl
Total Plasmid-Safe master-mix	12	

54| Add 12 µl of ATP-dependent DNase master-mix to each circularized DNA sample.

Component	Volume (µl)	Final concentration
Plasmid-Safe master-mix	12	
Circularized DNA	38	variable
Total	50	

55| Incubate in a thermocycler at 37 °C for 1 h, 70 °C for 30 min, hold at 4 °C.

56| Add 1× volumes (50 µl) of AMPure XP beads to the Plasmid-Safe ATP-Dependent treated DNA and purify as described in Step 12. Elute in 15 µl of TE pH 8.0. Transfer the supernatant to a new plate.

57| Pool DNA without beads and quantify by Qubit dsDNA HS assay and proper Qubit assay tubes.

■ **PAUSE POINT** Circularized DNA can be stored at −20 °C for several months.

***In vitro* cleavage of enzymatically purified, circularized gDNA ● TIMING 2 h**

58| *In vitro* cleavage with Cas9 and sgRNA. Dilute Cas9 and sgRNA as described in Steps 13 and 14, respectively. Setup *in vitro* cleavage master-mix:

Component	Volume (µl)	Final concentration
Cas9 Nuclease Reaction Buffer (10×)	5	1×
Cas9 Nuclease, <i>S. pyogenes</i> (1 µM)	4.5	90 nM
<i>In vitro</i> transcribed sgRNA (3 µM)	1.5	90 nM
Total cleavage master-mix	11	

▲ **CRITICAL STEP** Use RNase-free technique to avoid degrading sgRNA.

59| Incubate at room temperature for 10 min.

60| Add circularized DNA, diluted to a total volume of:

Component	Volume (μl)	Final concentration
Cleavage master-mix	11	
Plasmid-Safe DNase Treated DNA (125 ng)	39	2.5 ng/μl
Total	50	

61| Incubate in a thermocycler at 37 °C for 1 h, hold at 4 °C.

NOTE: A negative control sample must be included, consisting of circularized DNA incubated with Cas9 buffer, without Cas9:sgRNA.

62| Add 1× volumes (50 μl) of AMPure XP beads to the *in vitro* cleaved DNA and purify DNA as previously described in Step 12. Elute in 42 μl of TE pH 8.0, keeping the beads in solution.

▲ **CRITICAL STEP** Keep the beads in solution for the next enzymatic step.

Next-generation sequencing library preparation ● TIMING 4-6 h

63| A-tailing. Setup the A-tailing master mix:

Component	Volume (μl)	Final concentration
Kapa A-tailing Buffer (10×)	5	1×
Kapa A-tailing Enzyme	3	
Total A-tailing master-mix	8	

64| Add 8 μl of A-tailing master-mix to each eluted DNA sample with beads.

Component	Volume (μl)	Final concentration
A-tailing master-mix	8	
Cleaved DNA/beads	42	variable
Total	50	

65| Incubate in a thermocycler at 30 °C for 30 min, hold at 4 °C.

66| Add 1.8× volumes of PEG and NaCl SPRI solution (90 µl) to the A-tailed DNA and purify DNA as described in Step 12. Elute in 25 µl of TE pH 8.0, keeping the beads in solution.

▲ CRITICAL STEP Keep the beads in solution for the next enzymatic step.

67| Adapter ligation. Setup the adapter ligation master-mix:

Component	Volume (µl)	Final concentration
Kapa Ligation Buffer (5×)	10	1×
Kapa DNA Ligase	5	
NEBNext Adapter for Illumina (15 µM)	10	3 µM
Total master-mix	25	

68| Add 25 µl of adapter ligation master-mix to each A-tailed DNA sample with beads.

Component	Volume (µl)	Final concentration
Adapter ligation master-mix	25	
A-tailed DNA/beads	25	<i>variable</i>
Total	50	

▲ CRITICAL STEP Make single-use aliquots of NEB adapters to avoid adapter dimer formation due to freeze-thaw hydrolysis of the 3' T'.

69| Incubate in a thermocycler at 20 °C for 1 h, hold at 4 °C.

70| Add 1× volumes (50 µl) of PEG and NaCl SPRI solution to the adapter-ligated DNA and purify DNA as described in Step 12. Elute in 47 µl of TE pH 8.0, keeping the beads in solution.

▲ CRITICAL STEP Keep the beads in solution for the next enzymatic step.

71| USER enzyme. Add 3 µl of USER enzyme, provided with NEBNext® Multiplex Oligos for Illumina® (Dual Index Primers Set 1) to the adapter ligated DNA with beads. Incubate at 37 °C for 15 min.

72| Add 0.7× volumes (35 µl) of PEG and NaCl SPRI solution to the USER Enzyme treated DNA and purify as previously described in Step 12. Elute in 20 µl of TE pH 8.0. Transfer the supernatant to a new semi-skirted PCR plate and quantify by Qubit dsDNA HS assay and proper Qubit assay tubes (usually about 2-5 ng/µl).

73| Optionally, perform DNA size selection with PippinHT, using the 1.5% PippinHT cassette, with the size range of 250 to 850 bp, before proceeding to PCR. The recovered samples can be used directly for the next step.

? TROUBLESHOOTING

74| PCR. Setup a PCR master-mix for adding dual-index barcodes:

Component	Volume (μl)	Final concentration
Nuclease-free water	5	
2× Kapa HiFi HotStart Ready Mix	25	1×
Total master-mix	30	
Diluted PCR master-mix	30	
NEBNext i5 Primer (10 μM)	5	1 μM
NEBNext i7 Primer (10 μM)	5	1 μM
Total PCR mix	40	

▲ CRITICAL STEP Ensure that dual-index barcode sequences combinations are unique for each sample. Ideally, each sample would have unique i5 and i7 barcodes. Perform the PCR thermocycling conditions as shown below.

75| Add 40 μl of PCR master-mix to each sample of purified, USER enzyme treated DNA (~20ng).

Component	Volume (μl)	Final concentration
PCR mix	40	
USER enzyme treated DNA (~20ng)	10	~0.4 ng/μl
Total	50	

76| Perform the PCR using the following thermocycling conditions:

Step	Temperature	Time	Cycles
Denaturation	98 °C	45 s	1
Denaturation	98 °C	15 s	20
Annealing	65 °C	30 s	20
Extension	72 °C	30 s	20
Extension	72 °C	1 min	1
Hold	4°C		1

▲ **CRITICAL STEP** To avoid environmental contamination with complete adapter-ligated DNA, after library amplification, this PCR purification should be performed in a post-PCR hood.

77| Add 0.7× volumes (35 µl) of Agencourt AMPure XP beads to the PCR and purify as described in Step 12. Elute in 30 µl of TE pH 8.0. Transfer the supernatant to a new plate.

▲ **CRITICAL STEP** Run the PCR on a QIAxcel instrument as a library quality control, to check for adapter dimer formation. If adapter-dimers remain, further size selection should be performed.

■ **PAUSE POINT** PCR products can be stored at −20 °C for several months.

Quantify CIRCLE-seq libraries by digital droplet PCR (ddPCR) ● TIMING 6 h

CRITICAL: Standard quantitative qPCR can be substituted for ddPCR for CIRCLE-seq library quantification with similar performance, as long as adapter dimers are minimized.

78| Make 1:10 serial dilutions of 50 µl from 10^{−1} to 10^{−8} dilution of each sample from the library (PCR), starting with 5 µl of DNA and 45 µl of nuclease-free TE, and mix well. !

CAUTION Make serial dilution in a post-PCR hood to avoid contaminating the laboratory environment with amplified libraries.

79| Make ddPCR master-mix stock solution as follows:

Component	1 reaction (µl)	Final Concentration
2× ddPCR Supermix for probes	11	1×
Probe oSQT1310 (100 µM) (Table 2)	0.055	250 nM
Probe oSQT1311 (100 µM) (Table 2)	0.055	250 nM
Primer oSQT1274 (100 µM) (Table 2)	0.099	450 nM
Primer oSQT1275 (100 µM) (Table 2)	0.099	450 nM
Nuclease-free water	6.292	
Total ddPCR mix	17.6	

80| Assay 3 different dilution factors for each sample (10^{−6}, 10^{−7} and 10^{−8} from the library) in duplicate (in 96-well plate). A non-template control (NTC) is required. Add 17.6 µl of ddPCR master-mix to each sample.

Component	Volume (µl)	Final concentration
ddPCR mix	17.6	
Sample (add nuclease-free water into the NTC well)	4.4	<i>variable</i>
Total	22	

81| Seal the plate and spin down.

82| Droplet generation, thermocycling, and analysis: this step is performed using a QX200™ Droplet Digital™ PCR System (Bio-Rad). Place a Dg8 cartridge (8-well) into a DG8 cartridge holder. Add 70 µl of Droplet generation oil for probes into the oil row in the 8-well cartridge. Transfer 20 µl of sample from the semi-skirted PCR plate into the sample row in the 8-well cartridge. Cover the cartridge with the DG8 gasket (red rubber cover). Place into the droplet generator and close. It will automatically run. Take the cartridge out when finished. Transfer 40 µl from the droplet row of the 8-well cartridge into a semi-skirted Twin.tec PCR Plate 96-well (Eppendorf).

▲ CRITICAL STEP Only this specific PCR plate will fit into the Bio-Rad ddPCR instrument.

▲ CRITICAL STEP All the pipetting must be done slowly to ensure no air bubbles are being introduced into the well.

83| Touch PX1 PCR plate sealer to heat up to 180 °C. Place the Pierceable Foil Heat Seal on the plate with the red line on top and put into the plate sealer. Press seal.

84| Run PCR in thermocycler with the following program:

Cycling step	Temperature	Time	Cycles
Enzyme activation	95 °C	10 min	1
Denaturation	94 °C	30 s	40
Annealing/extension	60 °C	1 min	40
Enzyme deactivation	98 °C	10 min	1
Hold	4°C		1

85| After PCR, open the QuantaSoft software. Select the wells to be read. Choose ABS as the experiment type and ddPCR Supermix for probes. Select Ch1 Unknown in the target 1 and Ch2 Unknown in target 2. Click “Apply”, OK. Place the plate into the Droplet Reader™ (requires Droplet Reader Oil for running). Run. Choose FAM/HEX as dye set.

86| Analyze ddPCR Result. Gate the double-positive droplet population based on the negative control. Multiply the average of duplicate values by the dilution factor and by the five-fold dilution factor of the ddPCR reaction, as follows: Total copies/pl = # * 5 * dilution factor.

87| Pool all the samples in one library at equimolar concentrations. 1× pooled library should be in a total volume of 5 µl, $\sim 4.5 \times 10^9$ molecules.

Next-generation sequencing ● TIMING 1 d

88| Denature the pooled library ($\sim 4.5 \times 10^9$ molecules) by adding 5 μ l of NaOH 0.2N and incubate at room temperature for 5 min. Then, add 990 μ l of Hyb buffer (supplied with MiSeq® Reagent Kit v3 (600 cycle)).

89| Prepare the Phix control V3 (PhiX Control V3 KIT) as follows: mix 2 μ l of 10 nM PhiX control with 3 μ l of Tris-HCl 10 mM + 0.1% Tween-20, denature with 5 μ l of NaOH 0.2N and incubate at room temperature for 5 min. Add 990 μ l of Hyb buffer, to generate 20 μ M PhiX. Then, make a 12.5 μ M PhiX dilution, by mixing 375 μ l of the 20 μ M PhiX with 225 μ l of Hyb buffer. Add 50 μ l of the 12.5 μ M Phix to the denatured library (supplementing with 5% volume of Phix).

90| Clean the Flow Cell (supplied with MiSeq® Reagent Kit v3 (600 cycle)) with ultra-pure water, dry with lens tissues, followed by cleaning with alcohol wipes and lens tissue.

91| Load and sequence library using a MiSeq 600-cycle v3 kit according to manufacturer's instructions using MiSeq system. Sequencing is performed with 150 bp paired-end reads and 8 bp dual-index reads. 3-5 million reads are required for each sample, allowing the multiplexing of up to 5-8 samples.

92| After sequencing, copy the demultiplexed output FASTQ files to a location accessible to CIRCLE-seq analysis pipeline (installation instructions below).

CIRCLE-seq data analysis ● TIMING 1-3 h

93| Install CIRCLE-seq software, and the following and dependencies:

Python 2.7: <https://www.python.org/downloads/>

Burrows-Wheeler Aligner (bwa): <https://sourceforge.net/projects/bio-bwa/files/>

Samtools: <https://sourceforge.net/projects/samtools/files/samtools/>

Reference genome: download the appropriate reference genome, for example, the hg38 genome can be obtained from: <http://hgdownload.cse.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz>. If a reference genome for the species of interest is not available, the CIRCLE-seq analytical pipeline can be run in a reference-free mode and this step can be omitted.

94| Download and install CIRCLE-seq:

```
git clone https://github.com/tsailabSJ/circleseq.git
```

```
cd circleseq
```

```
pip install -r requirements.txt
```

95| Create a manifest file (.yaml). An example manifest is provided below; this example can be used with the example dataset included in the CIRCLE-seq software to test the software pipeline.

```
reference_genome: data/input/CIRCLEseq_test_genome.fa
```

```
analysis_folder: data/output

bwa: bwa

samtools: samtools

read_threshold: 4

window_size: 3

mapq_threshold: 50

start_threshold: 1

gap_threshold: 3

mismatch_threshold: 6

merged_analysis: True

samples:

  U2OS_EMX1:

    target: GAGTCCGAGCAGAAGAAGAANGG

    read1: data/input/EMX1.r1.fastq.gz

    read2: data/input/EMX1.r2.fastq.gz

    controlread1: data/input/EMX1_control.r1.fastq.gz

    controlread2: data/input/EMX1_control.r2.fastq.gz

    description: U2OS
```

Specify the *reference genome FASTA file*, *analysis output folder* and path to the *bwa* and *samtools* commands. Specify the target sequences and demultiplexed FASTQ files for both the nuclease and control samples. Multiple experiments can be analyzed together in a batch mode by listing them all in a single manifest file. CIRCLE-seq data can be analyzed either with or without a reference genome.

96 For standard reference-based analyses, run the command:

```
python /path/to/circleseq.py all --manifest /path/to/manifest.yaml
```

Alternately, for standard non-reference-based analyses, run the command:

```
python/path/to/circleseq.py reference-free --manifest /path/to/
manifest.yaml
```

97| When running the full pipeline, the results of each step are output to the *output_folder* in a separate folder for each step. The output folders and their respective contents are as follows:

output_folder	content
output_folder/aligned	Contains an alignment <i>.sam</i> , alignment <i>.bam</i> , sorted <i>.bam</i> , and <i>.bai</i> index file for each sample.
output_folder/fastq	Merged <i>.fastq.gz</i> files for each sample.
output_folder/identified	Contains tab-delimited <i>.txt</i> files for each sample containing the identified DSBs, control DSBs, filtered DSBs, and read quantification.
output_folder/visualization	Contains an <i>.svg</i> vector image representing an alignment of all detected off-targets to the target-site for each sample.

● TIMING

Step 1, cell culture: 1 week

Steps 2 - 8, genomic DNA isolation: 1 d

Step 9, sgRNA cloning into IVT plasmid, sequencing, *in vitro* transcription: 1 w

Steps 10 - 21, PCR, *in vitro* cleavage of PCR amplicon: 4 h

Steps 22 - 40, gDNA shearing, end repair, A tailing, adapter ligation: 1 d

Steps 58 - 57, Exonuclease, USER enzyme, intramolecular circularization and Plasmid-Safe ATP dependent DNase: 2 d

Steps 58 - 77, *in vitro* cleavage, A tailing, adapter ligation, USER enzyme, PCR: 1 d

Steps 78 - 92, ddPCR and NGS: 2 d

Steps 93 - 97, CIRCLE-seq data analysis: 1-3 h

ANTICIPATED RESULTS

Before proceeding with the full CIRCLE-seq protocol, it is important to first functionally test the ability of an RNP complex of interest to cleave a target-site containing DNA amplicon. It is anticipated that this cleavage reaction should proceed to near completion (see Fig. 3a).

Starting from high-quality genomic DNA, after random shearing and Agencourt AM-Pure XP bead purification (Steps 22-26), the sheared DNA fragments should have a median size of around 300 bp (See Fig. 3b), and a recovery after purification of approximately 75%.

Each CIRCLE-seq reaction requires approximately 25 µg of starting genomic DNA, as the expected yield of circularized DNA is approximately 1-2%.

After *in vitro* cleavage of the circularized DNA, the number of PCR cycles for enrichment is limited to 20 cycles, because higher numbers of cycles may increase the rate of PCR duplication. PCR enrichment of nuclease-cleaved fragments can be competitively suppressed by the presence of adapter dimers, which can be avoided by preamplification size selection (See Fig. 3c and 3d).

After a successful CIRCLE-seq experiment, the CIRCLE-seq analytical pipeline should output a range of off-target sites genome wide (see Fig. 4a–d), which are good candidates for further targeted sequencing validation in cells. For most targets, the number of raw on-target CIRCLE-seq read counts should be greater than 500, using 3-5 million paired-end sequence reads.

Acknowledgements

We thank N. Malinin for helpful comments and suggestions on this manuscript. This work was supported by St. Jude Children's Research Hospital and ALSAC, a National Institutes of Health (NIH) Director's Pioneer Award (DPI GM105378), NIH R35 GM118158, and NIH R01 GM107427 (to J.K.J.); and the Desmond and Ann Heathwood MGH Research Scholar Award (to J.K.J.).

References

1. Silva G et al. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr. Gene Ther* 11, 11–27 (2011). [PubMed: 21182466]
2. Urnov FD, Rebar EJ, Holmes MC, Zhang HS & Gregory PD Genome editing with engineered zinc finger nucleases. *Nature* 11, 636–646 (2010).
3. Cermak T et al. Efficient design and assembly of custom TALEN and other TAL effector based constructs for DNA targeting. *Nucleic Acids Res* 39 (12) (2011).
4. Sander JD & Joung JK CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol* 32, 347–355 (2014). [PubMed: 24584096]
5. Doudna JA & Charpentier E Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096 (2014). [PubMed: 25430774]
6. Rouet P, Smih F & Jasin M Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol* 14, 8096–8106 (1994). [PubMed: 7969147]
7. Fu Y et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol* 31, 822–826 (2013). [PubMed: 23792628]
8. Hsu PD et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol* 31, 827–832 (2013). [PubMed: 23873081]
9. Tsai SQ et al. CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* 14, 607–614 (2017). [PubMed: 28459458]
10. Tsai SQ & Joung JK Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nat. Rev. Genet* 17, 300 (2016). [PubMed: 27087594]
11. Wang X et al. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat. Biotechnol* 33, 175–178 (2015). [PubMed: 25599175]
12. Tsai SQ et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol* 33, 187–197 (2015). [PubMed: 25513782]
13. Kleinstiver BP et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol* 34, 869–874 (2016). [PubMed: 27347757]

14. Gabriel R et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol* 29, 816–823 (2011). [PubMed: 21822255]
15. Frock RL et al. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol* 33, 179–186 (2015). [PubMed: 25503383]
16. Hu J et al. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc* 11, 853–871 (2016). [PubMed: 27031497]
17. Crosetto N et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10, 361–365 (2013). [PubMed: 23503052]
18. Yan WX et al. BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun* 8, 15058 doi: 10.1038/ncomms15058 (2017). [PubMed: 28497783]
19. Leansing SV et al. DSBapture: *in situ* capture and sequencing of DNA breaks. *Nat. Methods* 13, 855–857 (2016). [PubMed: 27525976]
20. Canela A et al. DNA breaks and end resection measured genome-wide by end-sequencing. *Molecular Cell* 63, 898–911 (2016). [PubMed: 27477910]
21. Kim D et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12, 237–243 (2015). [PubMed: 25664545]
22. Kim D et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol* 34, 863–868 (2016). [PubMed: 27272384]
23. Cameron P et al. Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods*, 14, 600–606 (2017). [PubMed: 28459459]
24. Fisher S et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 12(1):R1 (2011). [PubMed: 21205303]
25. Robin JD et al. Comparison of DNA quantification methods for next generation sequencing. *Sci. Rep* 6:24067 (2016). [PubMed: 27048884]
26. Fu Y et al. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol* 32(3), 279–284 (2014). [PubMed: 24463574]

Editor Summary

This protocol describes CIRCLE-seq (Circularization for *in vitro* reporting of cleavage effects by sequencing), a sensitive and unbiased method for defining the on-target and off-target activity of CRISPR-Cas9 nucleases genome-wide.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1 |.

Overview of CIRCLE-seq workflow. (i-ii) Genomic DNA is isolated from cells (Steps 1-8). (iii) sgRNA is cloned into an *in vitro* transcription plasmid and *in vitro* transcribed (Step 9). (iv) Cas9:sgRNA RNP complex is first functionally tested for its ability to functionally cleave a PCR amplicon containing the intended target site to near completion (Steps 10-21). (v-vi) gDNA is sheared in an average length of 300 bp and circularized by intramolecular ligation in a procedure outlined in detail in Figure 2. Remaining uncircularized (linear) DNA molecules are degraded with exonuclease treatment (Steps 22-57). (vii) Circular DNA

molecules containing on- and off-target sites can subsequently be linearized with Cas9, releasing newly cleaved DNA ends (Steps 58-62) for (viii) adapter ligation, PCR amplification, and paired-end high-throughput sequencing (Steps 63-97). (Figure adapted with permission from Tsai et al, 2017⁹.)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2 |.

Detailed schematic overview of CIRCLE-seq method. Genomic DNA is randomly sheared to an average of ~300 bp, end-repaired, A-tailed, and ligated to uracil-containing stem-looped adapters. DNA molecules covalently closed with stem-looped adapters ligated to both ends are selected by treatment with a mixture of Lambda exonuclease I and *E. coli* exonuclease I, where partially ligated DNA with free 5' phosphate are degraded. The adapter-ligated DNA is treated with USER enzyme and T4 PNK, releasing 4 bp overhangs. DNA molecules are circularized at low concentrations, favoring intramolecular ligation.

Unwanted linear DNA is degraded with Plasmid-Safe ATP-dependent DNase. Circular DNA is cleaved with Cas9:sgRNA complex. Linearized DNA is ligated to sequencing adapters and amplified for high-throughput sequencing (Figure adapted with permission from Tsai et al, 2017⁹).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3 |. Quality control of CIRCLE-seq library preparation.

(a) Electropherogram showing *in vitro* cleavage of PCR amplicon containing the intended target site by Cas9:sgRNA RNP complex (red) versus an untreated control PCR product (blue). The red arrows indicate the expected cleavage products. Accuracy limitations of capillary electrophoretic size determination of higher molecular weight products explains the slight difference between the expected and observed sizes of the uncleaved PCR product. (b) QIAxcel electropherogram showing distribution of sheared genomic DNA; the median size of the fragmented DNA is 300 bp. (c) Example of electropherogram of suboptimal PCR for CIRCLE-seq library enrichment with high adapter-dimer percentage and (d) Example of electropherogram of optimal CIRCLE-seq library amplified after size-selection of adapter-ligated DNA fragments. DNA size-selection removes adapter dimers that compete with the adapter-ligated DNA fragments during PCR cycles.

Figure 4 | Expected results of CIRCLE-seq experiments.
(a, b) Visualization of detected off-targets identified by CIRCLE-seq aligned against intended target site for Cas9:sgRNA complexes targeted against *EMX1* and *VEGFA* genes, respectively. The intended target sequence is shown in the top line and off-target sites are ordered top to bottom by CIRCLE-seq read count. Mismatches to the intended target sequence are indicated by colored nucleotides, and read counts shown at the end of each line. The on-target site is marked with a star. For space considerations, this figure shows a listing of only the top CIRCLE-seq off-target sites detected in this experiment. (c, d) Manhattan plots of CIRCLE-seq detected off-target sites, with bar heights representing read count and organized by chromosomal position for *EMX1* and *VEGFA* site 1, respectively.

TABLE 1|

Comparison of methods for defining genome-wide CRISPR-Cas9 specificity

Method	Description	Advantages	Limitations
IDLV capture ^{11,10}	Cell-based method where integrase-defective lentiviral vectors are integrated with a selective marker into sites of nuclease-induced DSBs. Vector integration sites are recovered by linear amplification-mediated PCR (LAM-PCR), followed by high-throughput sequencing.	Certain cell types may be more amenable to infection with IDLV than transfection with dsODN tag.	Relatively insensitive due to the integration efficiency and the requirement of positive selection to overcome this; high level of background as IDLVs still retain some capability to randomly integrate into the cellular genome in the absence of nuclease-induced DSBs; IDLVs integrations can occur some distance from the nuclease-induced break so it may be more challenging to map sites.
GUIDE-seq ^{12,13,10}	Based on efficient integration of double stranded oligodeoxynucleotide (dsODN) tags at DSBs by NHEJ in living cells, followed by tag-specific amplification and high-throughput sequencing.	High efficiency of dsODN integration into DSBs enhances sensitivity; quantitative correlation between the numbers of GUIDE-seq reads with mutation frequencies in living cells.	Requires efficient cellular transfection of the dsODN tag, which can be challenging in sensitive cell types or <i>in vivo</i> settings.
HTGTS ^{15,16,10}	Detects off-target nuclease-induced DSBs by observation of translocation junctions between two nuclease-induced DSBs.	Can be applied to discovering nuclease-induced off-targets where the nucleases are delivered <i>in vivo</i> .	Nuclease-induced translocations are rare; translocations occur more frequently with sites in the same chromosome or in close nuclear proximity.
BLESS/BLISS ^{17,18,10} , End-seq ²⁰ , DSBCapture ¹⁹	Based on <i>in situ</i> ligation of adapters to transient nuclease-induced DSBs in fixed cells.	Do not require delivery and incorporation of exogenous DNA for detection	Lack information about nuclease-induced DSBs that were previously repaired by the cell repair machinery.
Digenome-seq ^{21,22,10}	<i>In vitro</i> method based on detection of Cas9-cleaved genomic DNA by whole genome sequencing.	Does not require PCR; has also been tested with base editors.	Does not enrich for nuclease-cleaved sequences and requires a large number of sequencing reads (~400 million); high level of background; yields only one-half of the cleaved site; lacks information about how cellular factors affect nuclease off-target activity.
Site-seq ²³	An <i>in vitro</i> method based on Cas9-cleavage of high molecular weight DNA, followed by enzymatic fragmentation, biotinylated adapter ligation, enrichment and sequencing.	Enriches for nuclease-cleaved fragments; reduces sequencing reads required.	Reads contain only one-half of the cleaved sites; lacks information about how cellular factors affects nuclease off-target activity.
CIRCLE-seq ⁹	<i>In vitro</i> method where genomic DNA is randomly fragmented, followed by circularization and generation of covalently closed dsDNA molecules. Circular dsDNAs are cleaved by Cas9 at on- and off-target sites, allowing the selective sequencing of nuclease-induced DSBs.	High enrichment so fewer reads required (3-5 million reads); reads contain both halves of the cleavage sites	Lacks information about how cellular factors affects nuclease off-target activity; requires large amount of gDNA.

TABLE 2 |

Primers, probes and adapter sequences

Primer	Sequence (5'-3')	Purpose
sgRNA-top	ATAGNNNNNNNNNNNNNNNNNN	Clone sgRNA into pCRL1 or equivalent vector (Step 9 (B))
sgRNA-bottom	AAACNNNNNNNNNNNNNNNNNN	Clone sgRNA into pCRL1 or equivalent vector (Step 9 (B))
M13 Reverse	CAGGAAACAGCTATGAC	Sequencing sgRNA cloned into pCRL1 (Step 9 (B, xii))
oSQT1143	GGAGCAGCTGGTCAGAGGGG	Amplify example <i>EMX1 locus</i> ²⁶ (Step 10)
oSQT1144	GGGAAGGGGGACACTGGGGA	Amplify example <i>EMX1 locus</i> ²⁶ (Step 10)
oSQT1288	/5Phos/CGGTGGACCGATGATC /ideoxyU/ ATCGGTCCACCG *T	CIRCLE-seq hairpin adapter (Step 27)
oSQT1274	AATGATACGCGACCAACCGAG	TruSeq F1 (Step 79)
oSQT1275	CAAGCAGAAGACGCATACGAGAT	TruSeq R1 (Step 79)
oSQT1310	/56-FAM/CCTACACGA/ZEN/CGCTCTTCCGATCT/3IABkFQ/	TruSeq probe (Step 79)
oSQT1311	/5HEX/TCGGAAGAG/ZEN/CACACGTCTGAACT/3IABkFQ/	TruSeq probe

* indicates phosphorothioate linkage.

/5Phos/ indicates 5' phosphorylation.

/ideoxyU/ indicates internal deoxyUridine

/56-FAM/ indicates 5' fluorescein modification

/5HEX/ / indicates 5' hexachlorofluorescein modification

/ZEN/ and /3IABkFQ/ are non-fluorescent quenchers.

TABLE 3 |**Troubleshooting**

Step	Problem	Possible reason	Possible solution
9 (B, ix)	Absence of colonies	Incompatible overhangs	Check that the overhangs in the sgRNA cloning oligonucleotides are correct
9 (B, ix)	Colonies growing on negative control plate	Incomplete digestions of pCRL1 or equivalent plasmid	Perform overnight digestion of the plasmid; increase amount of <i>Bsa</i> I; treat the digested plasmid with phosphatase to reduce self-ligation
9 (B, xii)	No sgRNA sequences	Incomplete digestion of cloning plasmid; ligation failure	Screen additional colonies; re-digest the plasmid; check the oligos sequences
21	No cleavage or low cleavage rate of PCR amplicon containing the intended target site	Impure input PCR product; sgRNA for the particular locus does not mediate cleavage; Cas9 lost activity	Purify PCR product; re-design sgRNA; check Cas9 activity with another amplicon and sgRNA
73	Poor library enrichment	Adapter dimers that compete with the adapter-ligated DNA during the PCR	Perform DNA size-selection of the adapter-ligated/USER treated DNA fragments using PippinHT