Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments

Wenjun Jiang & Ting F Zhu

School of Life Sciences, Center for Synthetic and Systems Biology, Ministry of Education Key Laboratory of Bioinformatics, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Tsinghua University, Beijing, China. Correspondence should be addressed to W.J. (jiangwj12@mails.tsinghua.edu.cn) or T.F.Z. (tzhu@biomed.tsinghua.edu.cn).

Published online 21 April 2016; doi:10.1038/nprot.2016.055

Cloning of long microbial genomic sequences is an essential tool in synthetic biology and genome engineering. Such long sequences are often difficult to obtain directly by traditional PCR or restriction enzyme digestion, and therefore the cloning of these sequences has remained a technical obstacle in molecular biology. Based on the *in vitro* application of RNA-guided Cas9 nuclease, the method of Cas9-assisted targeting of chromosome segments (CATCH) cleaves target DNA *in vitro* from intact bacterial chromosomes embedded in agarose plugs, which can be subsequently ligated with cloning vector through Gibson assembly. Here we describe an optimized protocol of CATCH cloning for the targeted cloning of long genomic sequences of up to 100 kb from microorganisms. The protocol uses standard laboratory equipment and takes ~8 h of bench time over several days, and it may potentially simplify and accelerate efforts to isolate and clone large gene clusters from microorganisms.

INTRODUCTION

The isolation and cloning of long, 100-kb genomic sequences have remained key technical challenges in molecular biology. Nevertheless, the ability to handle such long sequences is instrumental to some of the most important molecular work, such as sequencing particular genome regions, cloning large genes for protein expression, and efforts to design and engineer microorganisms[1–5]. The most commonly used method for cloning is to amplify the sequence of interest by PCR. However, traditional cloning methods based on PCR are often limited by the lengths (usually <10 kb) of the PCR products. One can also obtain long genomic sequences of interest by restriction enzyme digestion of total genomic DNA, and the methods for constructing long-insert bacterial artificial chromosome (BAC) libraries for the subsequent selection of desired sequences have been described in detail[6]. Yet the disadvantage here is that selecting a specific sequence of interest from a large number of restriction digest fragments can be very difficult and time-consuming[6]. In recent years, several alternative techniques such as single-strand overlapping annealing[7] and transformation-associated recombination (TAR)[8,9], have been developed to clone large bacterial gene clusters. However, the requirement of suitable restriction sites next to the target sequences has limited their versatility. Another route for obtaining long genomic sequences for cloning is the assembly of multiple short sequences such as synthetic DNA fragments or overlapping PCR products. Nonetheless, such methods often involve several assembly steps, and they tend to be labor-intensive and expensive[10,11].

The RNA-guided CRISPR/Cas9 endonuclease system provides better targeting specificity and versatility than traditional restriction enzymes (whose fixed recognition sites are typically limited to 6–8 bp)[12], which has led to the extensive development of Cas9-based genome editing *in vivo*[13–15]. Lately, the Cas9 system has been used *in vitro* for testing the enzyme's cleavage efficiency and for genotyping[16,17]. Notably, a recently developed method based on the in-solution Cas9 digestion of purified human genomic DNA and TAR allowed for the cloning of a long *NBS1* (*NBN*) gene into a yeast artificial chromosome[18]. We recently reported another new technique, CATCH[19], for the isolation and cloning of 100-kb bacterial genomic sequences by in-gel digestion of genomic DNA and subsequent Gibson assembly. The bacterial chromosomes are digested by RNA-guided Cas9 at designated target sites in low-melting-temperature agarose gel after cell lysis. The cloning vector, which shares 30-bp terminal-sequence overlaps with the target DNA at both ends, is ligated to the target DNA in a Gibson assembly mix[11], and the recombinant plasmid is electrotransformed into a cloning host (for experimental design and setup, see **Figs. 1** and **2**). By taking advantage of the programmable Cas9 nuclease system, CATCH allows the targeted isolation and cloning of long bacterial genomic sequences to be accomplished in a single step, thus potentially enabling many applications in biomedical research and the biotech industry.

Compared with other available methods, the CATCH in-gel cleavage method protects chromosomal DNA from shearing, which results in the successful cloning of 100-kb DNA sequences (a length sufficient for handling most of the known microbial gene clusters). It also allows for pulsed-field gel electrophoresis (PFGE) analysis after the RNA-guided Cas9 genome digestion, thus making the immediate troubleshooting and optimization assessment of the cleavage efficiency and off-target effects feasible. The current limitations of CATCH cloning include the requirement for a certain level of knowledge of the CRISPR/Cas9 system to design guide RNA pairs for each target sequence to cleave, the preparation of guide RNA pairs that involves *in vitro* transcription (IVT) of DNA templates prepared by overlapping PCR of multiple DNA oligonucleotides, and the preparation of a vector backbone with specific overhangs for each DNA segment to clone by PCR. In addition, intrinsic properties such as the off-target effects of the Cas9 system may affect the cleavage and cloning efficiency, especially when applied to more complex genome systems. Thus adaptations of the current methodology, such as combining the

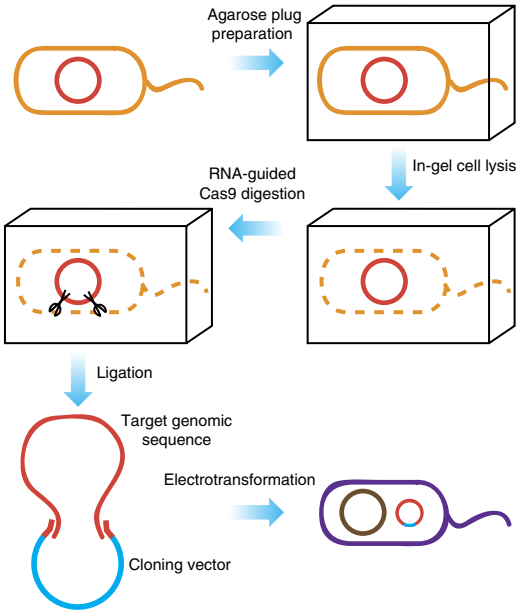

Figure 1 | Key workflow for CATCH cloning. The key workflow includes agarose plug preparation (Steps 12–15), in-gel cell lysis (Steps 16–19), in-gel RNA-guided Cas9 digestion (Steps 20–28), ligation of target genomic sequence with cloning vector (Step 36), and electrotransformation (Steps 37 and 38).

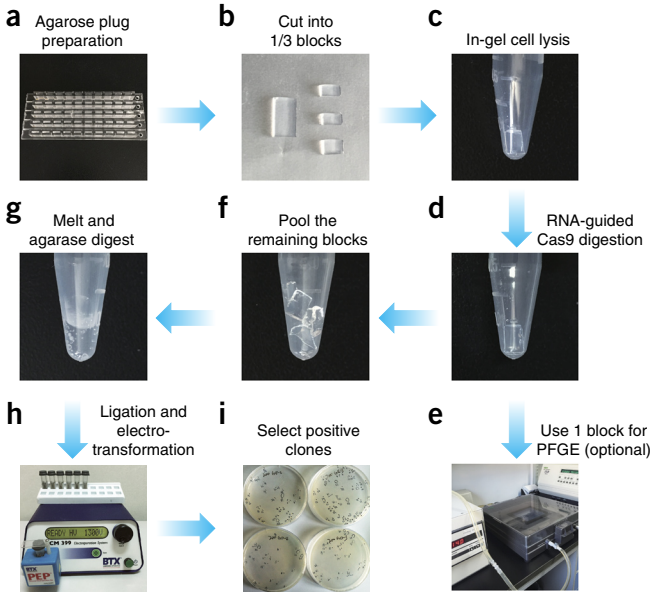

Figure 2 | Experimental setup for preparing agarose plugs, PFGE, and in-gel digestion. (**a**) The bacterial cells are embedded in low-melting-temperature agarose plugs using a 50-well disposable plug mold. (**b**) The agarose plugs are cut into three equal-sized blocks. (**c**) The bacterial cells in an agarose block (i.e., 1/3 of a plug) are lysed by lysozyme and proteinase K, leaving behind the intact chromosomes. (**d**) The agarose block should appear more transparent after the cell lysis compared with the one in **c**. RNA-guided Cas9 digestion cleaves the genomic DNA in a block. (**e**) (Optional) Use a block for PFGE to assess the cleavage efficiency and off-target effect. (**f**) The remainder of the cut blocks are pooled together in a 1.5-ml microcentrifuge tube. (**g**) The blocks are melted and digested by agarase. (**h**) The digested DNA is recovered, ligated to cloning vectors, and electrotransformed into host cells. (**i**) Colonies obtained from selective LB agar plates are PCR validated to identify positive clones.

in-gel digestion approach of CATCH with TAR[18], as well as the application of alternative RNA-guided nuclease systems[20,21], may facilitate efforts to clone large gene clusters from bacteria, as well as long genome segments from higher organisms. The ability to isolate intact, long genomic sequences by PFGE may also allow CATCH to be applied for sequencing purposes. Combined with the latest sequencing technologies, such as single-cell sequencing[22], optical mapping of long DNA segments[23], and methylation analysis[22,24,25], it may enable many more applications such as targeted sequencing of DNA methylation and targeted cloning and sequencing of large, disease-related genes in higher organisms.

Experimental design

Single guide RNA template preparation and *in vitro* transcription (Steps 1–11). The first step in CATCH cloning is to design effective single guide RNA (sgRNA) pairs to cleave the target genome segments from the source organism. Ideally, the whole genome sequence of the source organism should be available, although a draft genome with gaps should also suffice, and comprehensive PFGE analysis, PCR validation, and sequence validation should be carried out. A minimal requirement for selecting the target sequence is to obtain the flanking sequences next to the sequence of interest. Once the flanking sequences have been determined, search for protospacer-adjacent motif (PAM) sequence 'NGG' in the flanking sequences and the 20-bp sequence before the PAM sequence as the target sequence. Because the use of 'G' as the starting nucleotide (e.g., +1 position in **Fig. 3a**) of a transcript is required for T7 RNA polymerase *in vitro* transcription, if no 'G' is available as the starting nucleotide of the target sequence one may opt to add an additional 'G' to the sgRNA, which should not affect the cleavage efficiency and off-target effects. We note that the orientation of sgRNA targeting does not seem to influence the cleavage and cloning efficiency (see Figs. 3 and 4 of ref. 19). As an alternative to the manual selection of Cas9 target sequences, one can also refer to the online Optimized CRISPR Design tool (http://crispr.mit.edu/) or other tools to help avoid potential off-target effects, especially for genomes more complex than those of microorganisms[26,27]. We also note that sgRNA design for more effective use of the CRISPR/Cas9 system or other, similar systems is an area of active research; for example, truncated or extended versions of sgRNAs may reduce the off-target effects of Cas9 nucleases[28,29]. Thus future, more effective strategies for guide RNA design can be applied as modifications to our current protocol. The template for the IVT of sgRNA, which comprises a T7 promoter sequence, a target sequence and an sgRNA constant sequence contained in the crRNA-tracrRNA chimera, is typically ~117 bp (**Fig. 3a**). To minimize costs and to avoid synthesizing too-long DNA oligos, we recommend preparing the templates by overlapping PCR of three oligos, including a (target-specific) oligo containing a T7 promoter and a target sequence and two (universal) oligos of an sgRNA constant sequence contained in the crRNA-tracrRNA chimera (**Fig. 3a**). The overlapping PCR product can be purified by phenol/chloroform extraction and isopropanol precipitation, and it can then be resuspended in RNase-free water. The IVT can be performed with T7 RNA polymerase, and the product (sgRNA) can be purified by phenol/chloroform extraction and isopropanol precipitation (which is preferred because it ensures the recovery yield of short RNA molecules and effectively avoids RNase contamination). Alternatively, users may choose to chemically

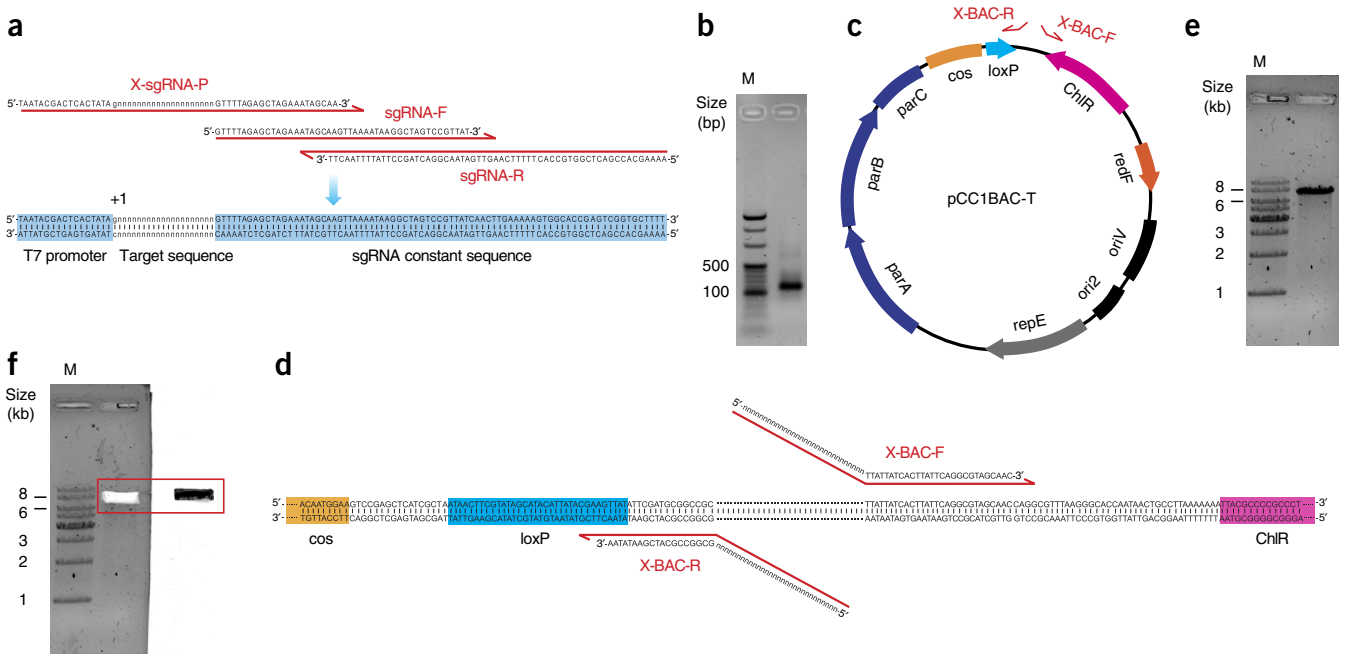

Figure 3 | Design and preparation of sgRNA and vector. (**a**) An example of an IVT template for sgRNA. The template is prepared by overlapping PCR of three oligos, including a target-specific oligo (X-sgRNA-P) containing the T7 promoter and target sequence, and two universal oligos (sgRNA-F and sgRNA-R) of an sgRNA constant sequence contained in the crRNA-tracrRNA chimera. The +1 position indicates the starting nucleotide for IVT. (**b**) Agarose gel electrophoresis of the overlapping PCR product: a full-length, 117-bp IVT template. (**c**,**d**) An example of the preparation of a BAC vector sharing 30-bp terminal-sequence overlaps with both ends of the target genomic segment. The PCR primers (X-BAC-F and X-BAC-R) contain an ~20-nt sequence that can anneal to the template plasmid and an ~30-nt overhang overlapping with the sequence at each end of the target DNA (indicated by multiple 'n's). The pCC1BAC-T plasmid may contain inserts of variable lengths, although longer (>50 kb) inserts help to minimize the contamination by template plasmids. (**e**,**f**) Conventional agarose gel electrophoresis to purify the ~7.5-kb BAC vectors from the long-insert template plasmids. The red box indicates the cutting area of the agarose gel to purify the BAC vectors. M, marker.

synthesize shorter versions of the guide RNAs[29] (for ease of synthesis), especially as the efficiency and cost of the chemical synthesis of RNA improve in the future.

Agarose plug preparation and in-gel cell lysis (Steps 12–19). CATCH cloning takes advantage of in-gel Cas9 cleavage, which protects the chromosomes from extensive shearing, permitting the isolation and cloning of intact, long DNA sequences (for experimental setup, see **Fig. 2**). The in-gel digestion of genomic DNA also allows for PFGE separation of the target sequences from the background genomic DNA and assessment of the cleavage efficiency and off-target effects, which makes the troubleshooting and optimization of the Cas9 cleavage more feasible. For preparation of the bacterial (e.g., *Escherichia coli* or *Bacillus subtilis*) agarose plugs for the subsequent in-gel Cas9 digestion, the bacterial cells are embedded in low-melting-temperature agarose plugs at a concentration of ~5×10^8 cells per ml. The plugs can be treated by lysozyme and proteinase K, and they can be successively washed with buffer according to the instructions for the Bio-Rad CHEF bacterial genomic DNA plug kit, leaving behind the genomic DNA. In the second wash, PMSF should be added to inactivate the remaining proteinase K, and 0.1× wash buffer should be used during the last wash for diluting the EDTA. Methods for making agarose plugs from other microorganisms such as the *Streptomyces* mycelia and yeast are comparable and have been described elsewhere[30,31]. The prepared genomic DNA agarose plugs can be stored in 1× wash buffer for up to 3 months at 4 °C, although another wash with 0.1× wash buffer should be carried out immediately before the Cas9 digestion. Note that when a new species or strain of microorganism is used for cloning, we recommend using PCR and Sanger sequencing to verify the target sequence in the genome of the source organism, as well as testing different cell concentrations for plug preparation for PFGE and cloning.

In-gel Cas9 digestion (Steps 20–28). The prepared agarose plugs are used for in-gel RNA-guided Cas9 digestion of the genomic DNA. For this, the agarose plugs should be first equilibrated in RNase-free cleavage buffer and transferred into a new tube containing the cleavage buffer and the Cas9 protein preassembled with sgRNA pairs, and then incubated at 37 °C for 2 h. After cleavage, the plugs should be washed with 0.1× wash buffer; 1/3 of an agarose plug can be used to evaluate the Cas9 cleavage efficiency and off-target effects by PFGE (optional; see **Box 1**), which are indicated by the presence of target and nontarget bands (see Anticipated Results and **Fig. 4**). The remaining plugs can be melted and subjected to digestion by agarase, followed by ethanol precipitation of the digested DNA and gentle resuspension in DNase-free water using wide-bore tips. Immediate ligation and transformation is recommended, although the DNA can be stored for several days at 4 °C. We have previously tested and optimized the conditions for in-gel RNA-guided Cas9 digestion, in an effort to understand the different factors that affect the cleavage efficiency[19] (see Supplementary Fig. 1 of ref. 19). The different conditions that we tested included agarose plug volume, bacterial cell concentration, Cas9 and sgRNA concentration, and incubation time.

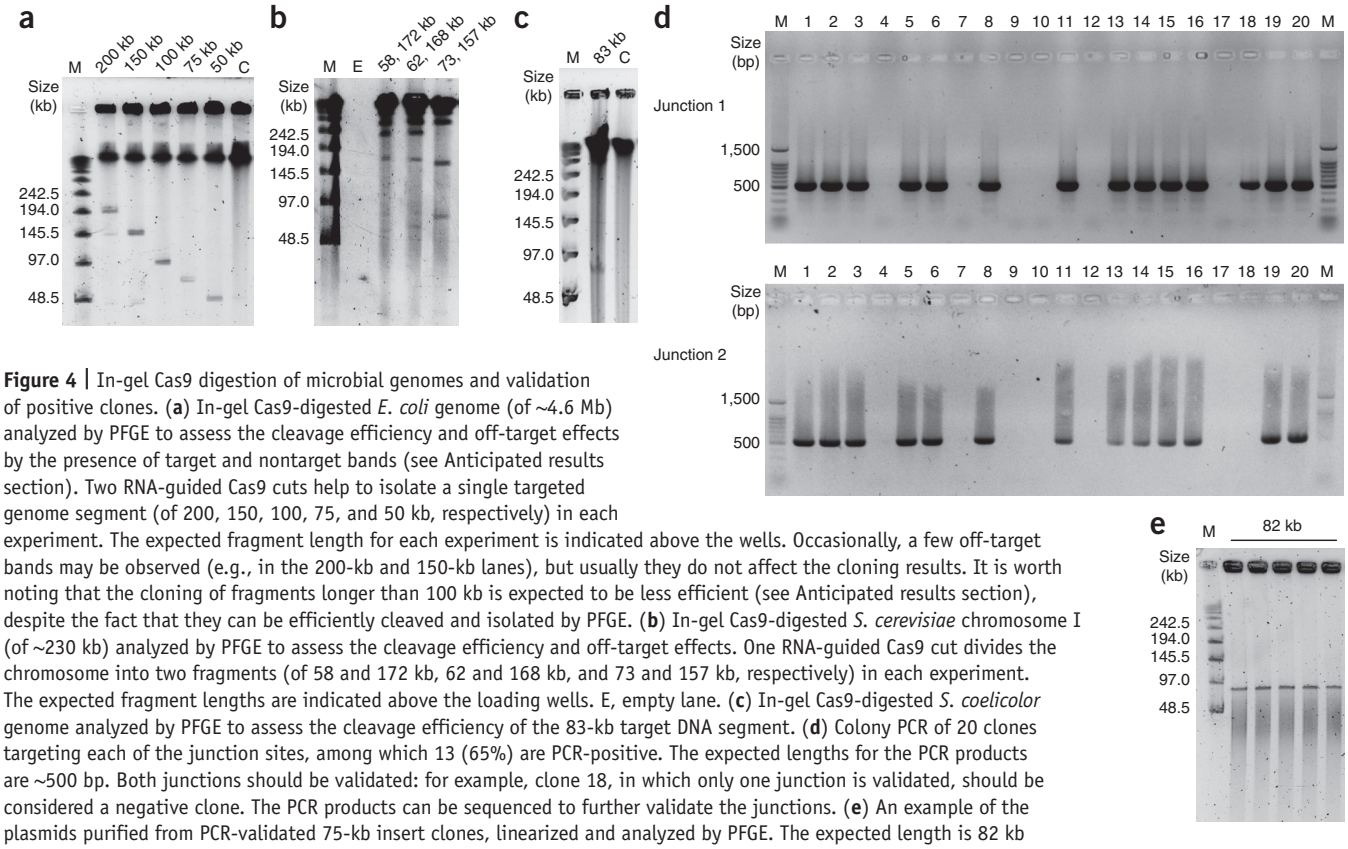

Figure 4 | In-gel Cas9 digestion of microbial genomes and validation of positive clones. (**a**) In-gel Cas9-digested *E. coli* genome (of ~4.6 Mb) analyzed by PFGE to assess the cleavage efficiency and off-target effects by the presence of target and nontarget bands (see Anticipated results section). Two RNA-guided Cas9 cuts help to isolate a single targeted genome segment (of 200, 150, 100, 75, and 50 kb, respectively) in each experiment. The expected fragment length for each experiment is indicated above the wells. Occasionally, a few off-target bands may be observed (e.g., in the 200-kb and 150-kb lanes), but usually they do not affect the cloning results. It is worth noting that the cloning of fragments longer than 100 kb is expected to be less efficient (see Anticipated results section), despite the fact that they can be efficiently cleaved and isolated by PFGE. (**b**) In-gel Cas9-digested *S. cerevisiae* chromosome I (of ~230 kb) analyzed by PFGE to assess the cleavage efficiency and off-target effects. One RNA-guided Cas9 cut divides the chromosome into two fragments (of 58 and 172 kb, 62 and 168 kb, and 73 and 157 kb, respectively) in each experiment. The expected fragment lengths are indicated above the loading wells. E, empty lane. (**c**) In-gel Cas9-digested *S. coelicolor* genome analyzed by PFGE to assess the cleavage efficiency of the 83-kb target DNA segment. (**d**) Colony PCR of 20 clones targeting each of the junction sites, among which 13 (65%) are PCR-positive. The expected lengths for the PCR products are ~500 bp. Both junctions should be validated: for example, clone 18, in which only one junction is validated, should be considered a negative clone. The PCR products can be sequenced to further validate the junctions. (**e**) An example of the plasmids purified from PCR-validated 75-kb insert clones, linearized and analyzed by PFGE. The expected length is 82 kb (insert plus vector sequence). C, control lane added with Cas9 but without sgRNA. M, marker.

We found that a bacterial cell concentration of ~5×10^8 per ml was ideal for the Cas9 cleavage and PFGE analysis (next section). We also found that final concentrations of 0.02–0.1 mg/ml Cas9 were optimal for the in-gel Cas9 digestion (note that the enzyme activity may vary depending on the expression and purification methods used to obtain the enzyme). Most important, using a final sgRNA concentration of >30 ng μl^{-1} seemed to be essential for the in-gel Cas9 cleavage, and thus preparation of high-quality sgRNA is crucial to the success of the CATCH experiment. Furthermore, we found that an incubation time of 1–2 h was sufficient for Cas9 diffusion and cleavage in agarose plugs.

PFGE analysis of cleaved DNA segments (optional). PFGE of the cleaved DNA segments is optional, as it is not required for cloning the target DNA sequences *per se*; however, performing PFGE is helpful for assessing the cleavage efficiency and off-target effects, as well as for troubleshooting (**Box 1**). In addition, PFGE can be used to isolate the cleaved long DNA fragments from the background genomic DNA, for other applications such as targeted DNA sequencing, in potential adaptations of the CATCH method. Approximately 1/3 of a Cas9-digested agarose plug is sufficient for PFGE analysis. The PFGE can be performed in 1% (wt/vol) Megabase agarose gel using the Bio-Rad CHEF Mapper XA system, and the gel can be stained by SYBR Gold (for its high sensitivity) for visualization using standard gel imaging systems. A clear band at the expected length should be observed (with the use of appropriate DNA markers, e.g., the Lambda Ladder PFG marker); the absence of clear bands suggests ineffective Cas9 cleavage, and the presence of nontarget bands suggests insufficient cleavage specificity by the RNA-guided Cas9 nuclease (troubleshooting guidelines can be seen in **Table 1**). In some cases, the plasmids carried by bacteria or the existence of DNase in the CHEF system may confound the PFGE analysis of cleaved DNA segments, and thus control experiments with Cas9 added but without sgRNAs are recommended (**Fig. 4**).

Preparation of cloning vectors (Steps 29–35). BAC vector, p15A vector, and other vectors can be used for cloning the target DNA sequence. BAC vectors can maintain insert sizes of up to a couple of 100 kb (ref. 6), whereas p15A vector is usually used with shorter inserts. The vectors prepared by PCR should share 30-bp terminal-sequence overlaps with the target DNA at both ends (for vector design, see **Fig. 3c**,**d**): each PCR primer contains an ~20-nt sequence that can anneal to the vector template and an ~30-nt overhang overlapping with the sequence at one end of the target DNA. Commercial plasmids containing pBAC or p15A vectors can be used as templates. Because contamination by the template plasmids may affect the cloning positive rates, the PCR products can be treated by DpnI endonuclease to remove the template plasmids, taking advantage of the ability of DpnI to cleave methylated plasmid DNA[32]. However, if achieving higher positive rates of cloning is crucial to the success of the experiment (e.g., for cloning long sequences of 50 kb or longer) and the use of DpnI cannot ensure the complete digestion of template plasmids, an improved method can be applied: use long-insert (>50 kb) plasmids as PCR templates, followed by agarose gel electrophoresis to separate and purify the vectors from the plasmid templates (**Fig. 3e**,**f**). The rationale here is that the long-insert (>50 kb) plasmids can be

Box 1 | PFGE analysis of cleaved DNA segments

1. After the in-gel Cas9 digestion, one of the six blocks (from Step 22) can be used to assess the cleavage efficiency and off-target effects by PFGE. Perform the PFGE with Lambda Ladder PFG marker in 1% Megabase agarose gel in 0.5× TBE buffer with the CHEF Mapper XA system set to auto-algorithm program with 5- to 250-kb parameters (6 V cm^{-1}, 0.22–21.79 s, 15 h 16 min, 120°, circulation at 14 °C).
2. After PFGE, stain the gel with SYBR Gold and visualize the DNA bands using a ChemiDoc XRS+ imaging system (or equivalent), which can be used to evaluate the Cas9 cleavage efficiency and off-target effects by the presence of target and nontarget bands (see Anticipated results section and **Fig. 4**).
? TROUBLESHOOTING

effectively separated from the vectors by conventional agarose gel electrophoresis, and thus contamination by template plasmids can be minimized. In our experience, the improved vector preparation method led to as much as twofold improvement of the positive rates. For users who are interested in applying the improved vector preparation method, the long-insert BAC plasmids can be either obtained by purifying the plasmids from prior CATCH clones or requested by contacting the authors of this protocol (see Reagents section). In our experience, negative clones may also arise from self-ligation of vectors at sites that share identical sequence of ≥6 bp (see Supplementary Fig. 3 in ref. 19), probably resulting from vectors nicked at sites with identical sequences, chewed by T5 5′–3′ exonuclease and self-ligated during Gibson assembly. Thus we recommend avoiding long identical sequences (≥6 bp) on the cloning vector.

Ligation and electrotransformation (Steps 36–38). The Cas9-cleaved target DNA can be recovered by melting of the Cas9-digested plugs followed by agarase digestion, ethanol precipitation, and resuspension in ultrapure water. Wide-bore tips should be used to avoid extensive shearing of the DNA. The recovered DNA typically has a viscous appearance and can be ligated to the corresponding cloning vector in a Gibson assembly mix containing T5 5′–3′ exonuclease, *Taq* DNA ligase and high-fidelity polymerase[11]. After the ligation, the mixture can be transformed into electrocompetent *E. coli* cells by electroporation. The electrocompetent *E. coli* cells can be either prepared (**Box 2**) or purchased (e.g., Epicentre TransforMax EPI300). The cells should recover at 37 °C for 2 h in 1 ml of SOC medium without antibiotics and then be plated on LB medium containing the corresponding antibiotics (depending on the antibiotic resistance gene on the vector; we used chloramphenicol) for incubation at 37 °C overnight. One should expect to obtain 50–100 (depending on the sizes of the inserts) colonies on the selective LB agar plates. For first-time users, we recommend performing a control experiment by transforming the pUC19 control DNA (which comes with the TransforMax EPI300 Electrocompetent *E. coli*) into the electrocompetent *E. coli* cells, to test the viability of the *E. coli* cells and the efficiency of the electrotransformation experiment.

Validation of positive clones (Steps 39–55). The colonies on selective LB agar plates can be validated by PCR and Sanger sequencing to target each of the junction sites (with one PCR primer on the vector and the other on the insert). We recommend picking ~20 colonies (for short inserts) or all of the colonies (for long inserts, which is important because of the lower cloning efficiency) for PCR validation (typically, there are 50–100 colonies on a plate). After colony PCR, the lengths of the PCR products can be examined by agarose electrophoresis (**Fig. 4d**). Colonies with correct PCR product lengths can be further validated by Sanger sequencing of the junction sequences (optional). Next, plasmids can be extracted from the cultured positive clones using the BACMAX DNA purification kit (for inserts >30 kb) or the QIAprep Miniprep kit (for inserts <30 kb). The purified plasmids can be linearized by λ-Terminase (for pCC1BAC vector) and analyzed by PFGE to validate the insert lengths (**Fig. 4e**).

Box 2 | Preparation of electrocompetent *E. coli* cells

Carry out the following steps in a decontaminated biosafety cabinet if possible; the use of surgical gloves and face masks is also recommended.

1. Culture the *E. coli* strain in 20 ml of LB No Salt medium (LB medium without NaCl added) overnight. Dilute 10 ml of the overnight *E. coli* culture into 500 ml of LB No Salt medium in a glass flask. Culture at 37 °C to OD 0.55–0.65 (typically in ~3 h) with shaking at 220 r.p.m.
2. Prepare in parallel an NaCl/ice/water bath by filling an ice bucket 1/3 full with water and 40 g of NaCl; add more ice until it reaches the top of water, and store the bucket in a cold room. Also prepare a dry ice/ethanol bath by mixing 95% (vol/vol) ethanol chilled to –20 °C with dry ice in a 1-l ice bucket.
3. Transfer the 500-ml *E. coli* culture into the NaCl/ice/water bath and chill it for 10 min. Equally divide the *E. coli* culture (250 ml each) into two large centrifuge bottles and centrifuge at 12,000*g* at 4 °C for 10 min. Discard the supernatant, resuspend each pellet with 50 ml of prechilled 10% (vol/vol) glycerol, and vortex briefly before transferring into prechilled 50-ml Falcon tubes. Centrifuge again at 12,000*g* at 4 °C for 10 min. Discard the supernatant, resuspend each pellet with 38 ml of prechilled 10% glycerol, and vortex briefly. Centrifuge again at 12,000*g* at 4 °C for 10 min. Discard the supernatant, resuspend each pellet in 25 ml of prechilled 10% glycerol, and vortex briefly. Centrifuge again at 12,000*g* at 4 °C for 10 min. Discard the supernatant and pipette out any residual fluid.
4. Resuspend each pellet in 500 µl of 10% glycerol, the final volume of which should be >1,000 µl. Divide the cells into aliquots in prechilled 1.5-ml microcentrifuge tubes (100 µl each), and immediately freeze them in the dry ice/ethanol bath. The prepared electrocompetent *E. coli* cells can be stored at –80 °C for up to 2 months until further use.

MATERIALS

REAGENTS

- BAC plasmid pCC1BAC-T (the authors can provide the long-insert BAC plasmid pCC1BAC-T for vector preparation for academic uses upon request)
- X-sgRNA-P primer, 10 μM (5′-TAATACGACTCACTATAgnnnnnnnnnnnnnnnnnnnnGTTTTAGAGCTAGAAATAGCAA-3′; the multiple 'n's indicate the 20-bp target sequence; Invitrogen)
- sgRNA-F primer, 10 μM (5′-GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC-3′; Invitrogen)
- sgRNA-R primer, 10 μM (5′-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT-3′; Invitrogen)
- X-BAC-F primer, 10 μM (5′-nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnTTATTATCACTTATTCAGGCGTAGCAAC-3′; the multiple 'n's indicate the 30-bp overlapping junction sequence; Invitrogen)
- X-BAC-R primer, 10 μM (5′-nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnGCGGCCGCATCGAATATAA-3′; the multiple 'n's indicate the 30-bp terminal sequence that overlaps with the target DNA; Invitrogen)
- BAC-vF primer, 10 μM (5′-AGTCCGAGCTCATCGCTAAT-3′; Invitrogen)
- BAC-vR primer, 10 μM (5′-TAAGGCAGTTATTGGTGCCCTT-3′; Invitrogen)
- X-vF primer, 10 μM (sequence designed based on the target sequence; Invitrogen)
- X-vR primer, 10 μM (sequence designed based on the target sequence; Invitrogen)
- Certified molecular biology agarose (Bio-Rad, cat. no. 1613100)
- Certified megabase agarose (Bio-Rad, cat. no. 1613109)
- 50× TAE buffer (Tris-acetate-EDTA; Thermo Scientific, cat. no. B49)
- 10× TBE buffer (Tris-borate-EDTA; Thermo Scientific, cat. no. B52)
- Autoclaved ddH_2O
- UltraPure DNase/RNase-free dH_2O (Life Technologies, cat. no. 10977) ▲ **CRITICAL** Although other sources of dH_2O may be used, in many steps of this protocol it is important to use DNase/RNase-free water.
- SYBR Green I nucleic acid gel stain (Invitrogen, cat. no. S-7563)
- SYBR Gold nucleic acid gel stain (Invitrogen, cat. no. S-11494)
- DTT, 1 M (Invitrogen, cat. no. P2325)
- 20-bp DNA Ladder (Takara, cat. no. 3420A)
- 100-bp DNA Ladder (Takara, cat. no. 3422A)
- 1-kb DNA Ladder (Takara, cat. no. 3426A)
- Lambda ladder PFG marker (NEB, cat. no. N0340S)
- 6× DNA loading buffer (Takara, cat. no. 9156)
- RNaseZap RNase decontamination solution (Ambion, cat. no. AM9780)
- Phenol:chloroform:isoamyl alcohol (25:24:1, pH >7.8; Solarbio, cat. no. P1012-100)
- Phenol:chloroform:isoamyl alcohol (25:24:1, pH <5.0; Solarbio, cat. no. P1011-100)
- HEPES, 1 M buffer solution (Gibco, cat. no. 15630-106)
- 3 M Sodium acetate, pH 5.5 (Ambion, cat. no. AM9740)
- RNase-free buffer kit (Ambion, cat. no. AM9010) ▲ **CRITICAL** It is crucial to use RNase-free buffer for all the steps involved in the preparation of sgRNA and Cas9 digestion.
- Q5 high-fidelity DNA polymerase (NEB, cat. no. M0491S)
- OneTaq quick-load 2× master mix (NEB, cat. no. M0486S)
- HiScribe T7 quick high yield RNA synthesis kit (NEB, cat. no. E2050S)
- Cas9 nuclease, *Streptococcus pyogenes* (NEB, cat. no. M0386M)
- Gibson assembly master mix (NEB, cat. no. E2611S)
- λ-Terminase (Epicentre, cat. no. LT44200)
- CHEF bacterial genomic DNA plug kit (Bio-Rad, cat. no. 1703592)
- CHEF yeast genomic DNA plug kit (Bio-Rad, cat. no. 1703593)
- Parafilm M (Parafilm, cat. no. PM996)
- PMSF protease inhibitor (Thermo Scientific, cat. no. 36978)
- GELase agarose gel-digesting preparation (Epicentre, cat. no. G09200)
- QIAquick gel extraction kit (Qiagen, cat. no. 28704)
- BACMAX DNA purification kit (Epicentre, cat. no. BMAX044)
- Absolute ethanol (e.g., Reagent Alcohol; Sigma-Aldrich, cat. no. 270741) ! **CAUTION** Absolute ethanol is highly flammable. Store it in a fireproof cabinet and use it with caution and away from open flame.
- Absolute isopropanol (e.g., 2-propanol; Sigma-Aldrich, cat. no. I9516-25ML) ! **CAUTION** Absolute isopropanol is highly flammable. Store it in a fireproof cabinet and use it with caution and away from open flame.
- TransforMax EPI300 Electrocompetent *E. coli*, including CopyControl Induction Solution and pUC19 control DNA (Epicentre, cat. no. EC300110)
- Chloramphenicol (Amresco, cat. no. 0230)
- Glycerol (Sigma-Aldrich, cat. no. G5516-100ML)
- SOC outgrowth medium (NEB, cat. no. B9020S)
- LB medium (Sigma-Aldrich, cat. no. L7275-100TAB)
- Microorganism (the microorganisms used in this work are model strains and are available at ATCC (http://www.atcc.org/), and their complete genome sequences are available at NCBI (http://www.ncbi.nlm.nih.gov/))

EQUIPMENT

- CHEF mapper XA chiller system (Bio-Rad, cat. no. 1703672)
- ECM 399 electroporator (BTX, cat. no. 45-0050INT)
- 1 mm Electroporation cuvette (BTX, cat. no. Bio-610)
- Pipettors (e.g., Research Plus pipettes, 0.1–2.5 μl, 0.5–10 μl, 2–20 μl, 20–200 μl, 100–1,000 μl; Eppendorf, cat. no. 3120000011, 3120000020, 3120000097, 3120000054, 3120000062)
- Clean pipette tips (e.g., epT.I.P.S. Racks pipette tips, 0.1–10 μl, 2–200 μl, 50–1,000 μl; Eppendorf, cat. no. 022491806, 022491831, 022491857)
- Wide-bore pipette tips (e.g., Axygen, cat. no. 383-T205WBCLRS)
- Thermal cycler (e.g., Veriti 96-well thermal cycler; Applied Biosystems, cat. no. 4375786)
- 1.5-ml Clean microcentrifuge tubes (e.g., 1.5-ml MAXYMum Recovery Snaplock Microcentrifuge tube; Corning, cat. no. MCT-150-L-C)
- 0.2-ml PCR tubes (e.g., 0.2-ml thin-wall PCR tubes with flat cap; Axygen, cat. no. PCR-02-C)
- Microwave oven
- Wide Mini ReadySub-Cell GT Cell (Bio-Rad, cat. no. 1704489EDU)
- PowerPac Universal Power Supply (Bio-Rad, cat. no. 1645070)
- ChemiDoc XRS+ System with Image La Software (Bio-Rad, cat. no. 1708265)
- UView mini transilluminator (Bio-Rad, cat. no. 1660531)
- Clean blade (e.g., scalpel blade sterile carbon steel no. 22; Swann-Morton, cat. no. 11792724)
- Biosafety cabinet (e.g., 1300 Series Class II, Type A2 biological safety cabinets; Fisher Scientific, cat. no. 13-261-339)
- Autoclave (e.g., Kinematica HMC vertical autoclave; Fisher Scientific, cat. no. 08-451-302)
- Ultra-low temperature freezer (e.g., Isotemp Basic −86 °C Ultra-Low Temperature Freezers; Fisher Scientific, cat. no. IE1386V)
- 4 °C Refrigerator (e.g., Isotemp General-Purpose Series Lab Refrigerators; Fisher Scientific, cat. no. 11670236)
- Precision balance (e.g., Mettler Toledo, cat. no. PL303-IC)
- Surgical glove (e.g., Cardinal Health Protexis latex classic surgical gloves; Fisher Scientific, cat. no. 19-163-104)
- Face mask (e.g., High Five Face Masks; Fisher Scientific, cat. no. 19-167-899)
- Microcentrifuge (e.g., Eppendorf Microcentrifuge 5424, cat. no. 022620428)
- Refrigerated centrifuge (e.g., Eppendorf Microcentrifuge 5424R, cat. no. 5404000537)
- NanoDrop 2000c spectrophotometer (Thermo Scientific, cat. no. ND-2000c)
- Incubator at 37 or 50 °C
- Water bath at 45, 50 or 70 °C
- Shaker at 37 °C, 250 r.p.m.
- Horizontal shaker at 80 r.p.m.

REAGENT SETUP

Ethanol solution for nucleic acid wash, 70% (vol/vol) Dilute absolute ethanol with ultrapure water at a 7:3 ratio. Freshly prepare and cool the solution to −20 °C before every use.

PMSF stock solution, 100 mM Dissolve 174 mg of the PMSF powder into 10 ml of absolute isopropanol. Divide the solution into aliquots in clean 1.5-ml microcentrifuge tubes, store them at room temperature (15–25 °C; up to 12 months), and protect them from light.

DTT stock solution, 10 mM Dilute 1 M DTT with ultrapure water at a 1:99 ratio, and store it at −20 °C (up to 6 months).

5× Cleavage buffer Mix the following solutions from RNase-free buffer kit. Divide the solution into aliquots in clean 1.5-ml microcentrifuge tubes and store them at −20 °C (up to 12 months).

Reagent	Volume (µl)	Final concentration (mM)
HEPES, pH 7.5 (1 M)	100	100
KCl (2 M)	375	750
EDTA, pH 8.0 (0.5 M)	1	0.5
$MgCl_2$ (1 M)	50	50
DEPC-treated water	474	
Total	1,000	

Chloramphenicol stock solution, 12.5 mg ml^{-1} Dissolve 0.125 g of chloramphenicol powder into 10 ml of absolute ethanol. Divide the solution into aliquots in clean 1.5-ml microcentrifuge tubes, store them at −20 °C (up to 12 months), and protect them from light.

Glycerol, 30% (vol/vol) Mix 30 ml of glycerol with 70 ml of ddH_2O, autoclave the mixture at 120 °C for 20 min, and store it at 4 °C (up to 6 months).

EQUIPMENT SETUP

Conventional agarose gel electrophoresis Prepare a sufficient amount of 1× TAE (Tris-acetate-EDTA) buffer diluted in ddH_2O from 50× TAE buffer stock for both agarose gel preparation and electrophoresis. To prepare a 3% or 0.8% (wt/vol) mini gel, add 0.9 g or 0.24 g of agarose and 30 ml of 1× TAE buffer into a 100-ml glass flask, mix it thoroughly and completely melt the agarose gel in a microwave oven. Cool the gel at room temperature for 10 min, add 0.3 µl of SYBR Green to the gel and mix it gently to avoid bubbles. Pour the gel into a mini gel mold with an appropriate comb and allow the gel to solidify. Load 5 µl of the DNA samples to be analyzed and the appropriate DNA markers, and run the gel at 8 V cm^{-1} for 15–40 min (depending on the experimental purposes). Photograph the gel with a ChemiDoc XRS+ Imaging System (or equivalent).

Pulse-field gel electrophoresis Prepare 2.5 l of 0.5× TBE (Tris-borate-EDTA) buffer diluted in ddH_2O from 10× TBE buffer stock for agarose gel preparation and electrophoresis. To prepare a 1% (wt/vol) medium-sized gel, add 1.2 g of Certified Megabase Agarose and 120 ml of 0.5× TBE buffer to a 250-ml glass flask, mix it thoroughly, and completely melt the agarose gel in a microwave oven. Cool the gel to ~60 °C, pour the gel into a PFGE gel-making mold with a 15-well comb, and let the gel solidify. Load the gel with agarose plugs (or 10 µl of liquid DNA sample of purified plasmids) and perform the PFGE in 0.5× TBE buffer with the CHEF Mapper XA System set to auto-algorithm program with 5–250 kb parameters (6 V cm^{-1}, 0.22–21.79 s, 15 h 16 min, 120°, circulation at 14 °C). After the run is completed, stain the gel with SYBR Gold solution (10 µl of SYBR Gold in 100 ml of 0.5× TBE buffer) for 15 min with shaking on a horizontal shaker gently, with a box cover to protect it from light. Photograph the gel with a ChemiDoc XRS+ Imaging System (or equivalent).

PROCEDURE

sgRNA template preparation and *in vitro* transcription ● TIMING 6 h

1| Perform overlapping PCR using one target-specific primer (X-sgRNA-P) and two universal primers (sgRNA-F and sgRNA-R) without any template, using the following program:

Reagent	Volume (µl)
5× Q5 reaction buffer	10
dNTPs (10 mM each)	1
X-sgRNA-P primer (10 µM)	2.5
sgRNA-F primer (10 µM)	1
sgRNA-R primer (10 µM)	2.5
Q5 High-fidelity DNA polymerase (2 U µl^{-1})	0.5
Ultrapure water	32.5
Total	50

	Cycles	Temperature (°C)	Time
Initial denaturation	1	98	30 s
Amplification	30	98	10 s
		55	20 s
		72	15 s
Final extension	1	72	2 min
Hold	1	4	

2| Mix 3 µl of the PCR product with 0.5 µl of 6× DNA loading buffer, and perform electrophoresis in 3% (wt/vol) mini agarose gel in 1× TAE buffer, along with 20-bp DNA marker to validate the length of the PCR product.

▲ CRITICAL STEP It is recommended that Steps 3–11 be carried out in a biosafety cabinet decontaminated with RNaseZAP to avoid RNase contamination. The use of surgical gloves and face masks is also recommended.

3| Adjust the volume of the PCR product to 180 µl in a clean 1.5-ml microcentrifuge tube by adding ~130 µl of ultrapure water. Add an equal volume (180 µl) of phenol:chloroform:isoamyl alcohol (25:24:1, pH >7.8) and vortex it vigorously for 15 s. Allow the homogenate to remain at room temperature (15–25 °C) for 3 min before centrifugation at 12,000*g* at 4 °C for 10 min.

4| Carefully transfer the upper aqueous layer (180 µl) into a clean 1.5-ml microcentrifuge tube without touching the interphase or the lower organic layer with the pipette tip. Add 20 µl (~10% of the final volume) of 3 M sodium acetate (pH 5.5) and mix it thoroughly.

5| Add an equal volume (200 µl) of isopropanol and mix it thoroughly. Incubate the mixture at –80° C for 10 min, and collect the pellet (precipitated DNA) by centrifugation at 12,000*g* at 4 °C for 30 min.

6| Carefully remove and discard the supernatant, rinse the pellet twice by adding 500 µl of freshly prepared ice-cold 70% ethanol and centrifuge it at 12,000*g* for 10 min at 4 °C before discarding the supernatant.

7| Centrifuge the microcentrifuge tube briefly and aspirate any residual ethanol using a 10-µl pipette tip. Air-dry the pellet and resuspend it in 30 µl of ultrapure water. Use 1 µl of the purified IVT template to measure the concentration and purity of DNA by NanoDrop. A concentration of between 30 and 100 ng μl^{-1} and A260/A280 >1.8 should be ideal.
■ **PAUSE POINT** The IVT template can be stored at –20 °C for up to 6 months.

8| Perform *in vitro* transcription by mixing the IVT template from Step 7 with the following reagents from the HiScribe T7 Quick High Yield RNA synthesis kit, and incubate the mixture at 37 °C for 2 h. The reaction generates large quantities of RNA, and thus the reaction mixture may appear viscous after the incubation.

Reagent	Volume (µl)
NTP buffer mix	16.5
Template DNA (30–100 ng μl^{-1})	30
T7 RNA polymerase mix	3.5
Total	50

9| To remove the IVT template, add 130 µl of ultrapure water to dilute the reaction. Then add 4 µl of DNase I (RNase-free; included in the HiScribe T7 Quick High Yield RNA synthesis kit), mix it thoroughly, and incubate at 37 °C for 15 min.

10| Re-perform Steps 3–7 to purify the *in vitro*–transcribed sgRNA, with the exception that phenol:chloroform:isoamyl alcohol (25:24:1, pH <5.0) should be used for the extraction, followed by isopropanol precipitation. Resuspend the pellet in 100 µl of ultrapure water, and use 1 µl to measure the concentration and purity of sgRNA by NanoDrop. A concentration of 300–1,000 ng μl^{-1} and A260/A280 >2.0 should be ideal.
▲ **CRITICAL STEP** The pH of phenol:chloroform:isoamyl alcohol used for DNA and RNA extraction is different (pH >7.8 for DNA and pH <5.0 for RNA).
? TROUBLESHOOTING

11| Dilute the prepared sgRNA to a concentration of 300 ng μl^{-1} with ultrapure water, and prepare 20-µl aliquots in individual 0.2-ml tubes and store them at –80° C.
▲ **CRITICAL STEP** Both the IVT template and the sgRNA are purified by phenol/chloroform extraction and isopropanol precipitation, as described above. Alternatively, one may use commercial kits to purify the nucleic acids, although the recovery rate may be lower and contamination by RNase may occur.
■ **PAUSE POINT** The prepared sgRNA can be stored at –80 °C for up to 3 months.

Agarose plug preparation and in-gel cell lysis ● TIMING 2 d
12| Inoculate the microbial culture into a 5-ml volume of the appropriate medium, and grow it with agitation to an optimized concentration at an appropriate temperature. For example, *E. coli* at a concentration of ~5 × 10^8 cells per ml in agarose plug is optimal for both the cell lysis and Cas9 digestion[19]. To culture *E. coli* for agarose plug preparation, inoculate 50 µl of the overnight culture in 5 ml of LB liquid medium and incubate it at 37 °C with shaking for 2–3 h, until the OD_{600} reaches 1.0 (~10^9 cells per ml).
▲ **CRITICAL STEP** For the microorganism used for CATCH cloning for the first time, different cell concentrations may need to be tested for optimal cleavage and cloning efficiency[19]. Use a hemocytometer under a microscope to estimate the cell concentration if necessary.

13| During the culture, melt 500 µl of 2% CleanCut agarose solution using a hot water bath at 70 °C and then cool the solution to 50 °C in a water bath.

14| Centrifuge 500 µl of the bacterial culture from Step 12 at 12,000*g* at 4 °C for 2 min in a clean 1.5-ml microcentrifuge tube, carefully discard the supernatant, and resuspend to 500 µl with cell suspension buffer.

15| Gently and thoroughly mix the 500-µl cell suspension with an equal volume (500 µl) of melted 2% CleanCut agarose from Step 13 using a 1-ml pipette tip, to give a final concentration of ~5×10^8 cells per ml in agarose plugs. Maintain the mixture at 50 °C and transfer the mixture to a 50-well disposable plug mold slowly using a 200-µl pipette. 1 ml of the cell–agarose mixture can fill 10 wells in the mold. Place the mold at 4 °C for 10 min to allow the agarose to solidify.
▲ **CRITICAL STEP** The 2% CleanCut agarose and the cell–agarose mixture can be sticky, especially when cooled. Thus, we recommend pipetting slowly to avoid air bubbles.

16| While waiting for the agarose plugs to solidify, prepare the lysozyme (for bacterial cells) or lyticase (for yeast cells) solution by adding 100 µl of the lysozyme stock into 2.5 ml of lysozyme buffer (or the same volume of lyticase stock to a lyticase buffer). Divide 83 µl of the solution into 30 individual clean 1.5-ml microcentrifuge tubes for immediate use.

17| Peel off the seal from the mold and push the solidified agarose plugs onto a clean sheet of Parafilm. Equally cut each plug into three blocks with a clean blade, and transfer each block into a 1.5-ml microcentrifuge tube containing the lysozyme or lyticase solution from Step 16. Make sure that every block is submerged in the solution, and incubate it at 37 °C for 2 h.

18| Remove the lysozyme or lyticase solution and rinse the plugs once with ultrapure water. Prepare the proteinase K solution by adding 100 µl of the proteinase K stock into 2.5 ml of proteinase K reaction buffer. Add 83 µl of the proteinase K solution into each tube containing the block. Make sure that every block is submerged in the solution and incubate at 50 °C without agitation overnight. The blocks can be incubated in proteinase K for up to 4 d if necessary for the complete lysis of different microbial cells.

19| After the proteinase K digestion, prepare 1× wash buffer by diluting 2.5 ml of 10× wash buffer stock into 22.5 ml of ultrapure water. Remove proteinase K solution and wash the plugs four times with 1× wash buffer for 1 h each at room temperature by gentle agitation (for each block, use 200 µl of 1× wash buffer during each wash). During the third round of washing, use 1× wash buffer containing 1 mM PMSF to inactivate the residual proteinase K. Prepare the 1× wash buffer containing 1 mM PMSF immediately before use by adding 60 µl of the 100 mM PMSF stock into 6 ml of 1× wash buffer. After all four rounds of washing, store the plugs at 4 °C in 1× wash buffer.
■ **PAUSE POINT** The prepared genomic DNA agarose plugs can be stored in 1× wash buffer at 4 °C for up to 3 months, although another wash with 0.1× wash buffer should be carried out immediately before Cas9 digestion.

In-gel Cas9 digestion ● TIMING 3 h

20| For the in-gel Cas9 cleavage of each target sequence, six agarose blocks are used. Perform another wash with 0.1× wash buffer (diluted from 1× wash buffer with ultrapure water) immediately before Cas9 digestion for diluting the EDTA. Discard the wash buffer and add 100 µl of 1× cleavage buffer (diluted from 5× cleavage buffer) into each tube, and equilibrate the agarose blocks at room temperature for 30 min.

21| During the equilibration, assemble Cas9 with sgRNAs by mixing the following reagents, and incubate the mixture at 37 °C for 20 min. A pair of sgRNAs (X1-sgRNA and X2-sgRNA, where 'X' stands for the name of the sequence of interest) prepared from Step 11 are used for the RNA-guided Cas9 cleavage at the two flanking sites of the sequence of interest.

Reagent	Volume (µl)
5× Cleavage buffer	60
DTT (10 mM)	15
X1-sgRNA (300 ng μl^{-1})	30
X2-sgRNA (300 ng μl^{-1})	30
Cas9 (20 µM)	3
RNase-free water	162
Total	300

22| Discard the cleavage buffer from the microcentrifuge tube in Step 20, add 50 µl of the prepared cleavage solution from Step 21 into each microcentrifuge tube, and incubate at 37 °C for 2 h, allowing the diffusion and digestion by Cas9 to occur in the agarose matrix.

23| After the incubation, discard the cleavage solution and pool all of the available blocks for each target sequence into a clean 1.5-ml microcentrifuge tube. Wash the plugs with 1 ml of 0.1× wash buffer, discard the wash buffer, and add 500 µl of 1× GELase buffer (diluted from 50× GELase buffer with ultrapure water). Incubate the mixture at room temperature for 1 h, and discard the 1× GELase buffer.

24| Thoroughly melt the agarose blocks by incubating them in a 70 °C water bath for 5 min. Flick the microcentrifuge tube gently to assess the consistency and appearance of the agarose. When the sample has turned into liquid, transfer the microcentrifuge tube containing the molten agarose into a 45 °C water bath and equilibrate for 10 min.
▲ **CRITICAL STEP** Incubation temperatures higher than 70 °C may denature DNA.

25| Add 1 µl of GELase enzyme slowly to the agarose. Stir the mixture with a 10-µl tip very gently to avoid shearing of the long DNA fragments, and incubate the mixture at 45 °C for 30 min.
▲ **CRITICAL STEP** The GELase is temperature sensitive, and thus incubation temperatures higher than 45 °C may inactivate the enzyme.

26| After GELase digestion, cool the sample to room temperature for 5 min, and flick the microcentrifuge tube gently to examine the appearance of the agarose. If the sample remains liquid, add an equal volume (~170 µl) of 5 M NH_4OAc to the digested agarose, and gently invert the microcentrifuge tube several times to mix thoroughly.

27| Add a 4× volume (~680 µl) of room-temperature ethanol to the microcentrifuge tube and gently invert the tube several times to mix thoroughly. Incubate the mixture at room temperature for 2 h, and pellet the DNA by centrifugation at 12,000*g* at room temperature for 15 min.

28| Carefully remove the supernatant, wash the pellet with 70% ethanol twice, and carefully discard the supernatant. Air-dry the pellet at room temperature for 5 min, add 20 µl of ultrapure water to the pellet, and place the sample at 4 °C for 2 h to dissolve the DNA.
▲ **CRITICAL STEP** Do not pipette the DNA, to avoid shearing of the long DNA fragments. After the DNA molecules are dissolved, the solution should appear viscous.
■ **PAUSE POINT** The obtained DNA can be stored at 4 °C for several days, although immediate ligation is recommended.

Preparation of cloning vectors ● **TIMING 4 h**
29| To prepare the BAC vector for ligation, perform PCR using primers with 30-nt overhangs using the following program:

Reagent	Volume (µl)
5× Q5 reaction buffer	10
dNTPs (10 mM each)	1
X-BAC-F primer (10 µM)	2.5
X-BAC-R primer (10 µM)	2.5
pCC1BAC-T (20 ng μl^{-1})	0.5
Q5 High-fidelity DNA polymerase (2 U μl^{-1})	0.5
ddH_2O	33
Total	50

	Cycles	Temperature (°C)	Time
Initial denaturation	1	98	30 s
Amplification	35	98	10 s
		65	20 s
		72	5 min
Final extension	1	72	10 min
Hold	1	4	

30| Add 10 µl of 6× DNA loading buffer to the PCR product, mix it thoroughly, and load the sample in a 0.8% agarose gel. Perform conventional electrophoresis in 1× TAE buffer along with a 1-kb DNA marker (see Equipment Setup section).

31| Excise the band (~7.5 kb) from the agarose gel with a clean blade using the UView Mini Transilluminator. Weigh the gel slice in a 1.5-ml microcentrifuge tube (typically <300 mg). For slices >300 mg in weight, cut the gel into smaller pieces and divide it into multiple tubes.

32| Add a 3× volume of buffer QG (included in the QIAquick gel extraction kit) to 1 volume of gel (100 mg of gel approximates 100 µl). Incubate the mixture at room temperature for ~10 min until the gel slice has completely dissolved. Mix by inverting the microcentrifuge tube gently several times during the incubation to help dissolve the agarose gel.
▲ **CRITICAL STEP** It is not recommended to heat the mixture for dissolving the gel.

33| Add a volume equal to that of the gel of isopropanol to the sample, mix it thoroughly, transfer the sample to a QIAquick column, and centrifuge it at 12,000*g* for 1 min. Discard the flow-through and add another 500 µl of buffer QG to the QIAquick column; then centrifuge at 12,000*g* for 1 min, and discard the flow-through.

34| Add 750 µl of buffer PE (with ethanol added, as indicated by the manufacturer) to the QIAquick column to wash it. Allow the QIAquick column to remain still for 2–5 min before centrifugation at 12,000*g* for 1 min. Discard the flow-through and centrifuge the QIAquick column for another 2 min at 12,000*g*.

35| Place the QIAquick column into a clean 1.5-ml microcentrifuge tube, and air-dry the column for 5 min. To elute DNA, add 20 µl of ultrapure water to the center of the QIAquick membrane. Allow the column to remain still for 5 min, and then centrifuge the column at 12,000*g* for 1 min and collect the eluted DNA. Use 1 µl to measure the concentration of the cloning vector using NanoDrop, which typically should be 30–50 ng μl^{-1}.
■ **PAUSE POINT** The prepared vectors can be stored at –20 °C for several weeks, although immediate ligation is recommended and repeated freezing-thawing should be avoided.

Ligation and electrotransformation ● TIMING 2 h

36| For ligation, gently mix the following reagents with a wide-bore tip in a 0.2-ml PCR tube, and incubate the sample in a thermal cycler at 50 °C for 1 h and hold it at 4 °C.

Reagent	Volume (µl)
2× Gibson assembly mix	5
Cas9 digested genomic DNA (from Step 28)	4
Cloning vector (30–50 ng μl^{-1} from Step 35)	1
Total	10

37| After the ligation, use 2 µl of the mixture for the electrotransformation into 50 µl of TransforMax EPI300 Electrocompetent *E. coli* cells or prepared electrocompetent *E. coli* cells (**Box 2**) in a 1-mm cuvette using the BTX ECM 399 Electroporation system set to 1,300 V electroporation voltage. The peak voltage should be no lower than 1,200 V, and the time constant should be between 4 and 6 ms. Add 500 µl of warm (~37 °C) SOC into the cuvette immediately. Resuspend the *E. coli* cells and transfer them into a clean 1.5-ml microcentrifuge tube. Add another 500 µl of warm SOC into the cuvette, pipette it twice to collect the remaining cells, and transfer the mixture into the same microcentrifuge tube. Incubate the microcentrifuge tube at 37 °C for 2 h with shaking at 200 r.p.m., which allows the cells to recover and to develop resistance to antibiotics.
▲ **CRITICAL STEP** The electrocompetent *E. coli* cells should be kept on ice before and after the electrotransformation until SOC is added.
? TROUBLESHOOTING

38| After incubation, plate the cells from the 1 ml of SOC culture (use 200 µl per plate) onto three or more LB agar plates containing 12.5 µg/ml chloramphenicol. Dry the plates in a sterile laminar flow hood or biosafety cabinet to allow the cells to be absorbed into the medium. Incubate the plates at 37 °C overnight.
? TROUBLESHOOTING

Validation of positive clones ● TIMING 1 d

39| Pick ~20 colonies from the selective LB agar plate from Step 38 with individual tips. Inoculate each of the picked colonies onto a new LB agar plate containing 12.5 μg/ml chloramphenicol and label them sequentially. Immediately after inoculation, dip each individual tip into a clean 0.2-ml tube containing 5 μl of water, to be used as template for colony PCR (Step 40). Incubate the plates at 37 °C for 4 h.

40| Perform two colony PCR reactions for each clone to target both of the ligation sites (for each ligation site, use one PCR primer on the vector and the other on the insert; see Experimental design section). Use the following program and examine the PCR products by conventional electrophoresis in 0.8% agarose gel with 100-bp DNA marker (see Equipment Setup section), and look for the PCR products of expected lengths.

? TROUBLESHOOTING

Reagent	Volume (μl) for each colony	Total volume (μl) for 20 colonies
OneTaq Quick-Load 2× Master Mix	5	100
BAC-vF primer (10 μM)	0.25	5
X-vR primer (10 μM)	0.25	5
Water the tip was dipped in (Step 39)	0.5	—
ddH_2O	4	80
Total	10	190

Reagent	Volume (μl) for each colony	Total volume (μl) for 20 colonies
OneTaq Quick-Load 2× Master Mix	5	100
X-vF primer (10 μM)	0.25	5
BAC-vR primer (10 μM)	0.25	5
Water the tip was dipped in (Step 39)	0.5	—
ddH_2O	4	80
Total	10	190

	Cycles	Temperature (°C)	Time
Initial denaturation	1	94	5 min
Amplification	30	94	20 s
		56	20 s
		68	30 s
Final extension	1	68	5 min
Hold	1	4	

41| Choose five of the PCR-validated clones and inoculate each into individual 5 ml of LB liquid medium containing 12.5 μg/ml chloramphenicol and 30 μl of CopyControl BAC autoinduction solution. Incubate the mixture at 37 °C with shaking at 200 r.p.m. overnight.

42| Transfer 500 μl of the overnight culture mixed with an equal volume of sterilized 30% glycerol to a clean 1.5-ml microcentrifuge tube and store it at −80 °C.

43| Use 1.5 ml of the remaining culture to centrifuge at 12,000*g* for 1 min to pellet the cells. Discard the supernatant, add 200 μl of chilled BACMAX solution 1 to the pellet, and vortex to resuspend the pellet.

44| Add 400 μl of BACMAX solution 2 to lyse the cells. Mix by gently inverting the microcentrifuge tube a few times to avoid shearing of the long-insert plasmids. The lysis time should be no longer than 5 min.

45| Add 300 µl of chilled BACMAX solution 3. Mix by gently inverting the microcentrifuge tube a few times. Incubate the mixture on ice for 15 min, and centrifuge it at 12,000*g* at 4 °C for 15 min.

46| Slowly aspirate the fluid below the meniscus and transfer it into a 1.5-ml microcentrifuge tube, avoiding the meniscus and the pellet.
▲ **CRITICAL STEP** Use wide-bore tips to minimize shearing of the long-insert plasmids.

47| Add a 0.6× volume (~540 µl) of room-temperature isopropanol. Mix by gently inverting the microcentrifuge tube a few times, and centrifuge at 12,000*g* at 4 °C for 15 min. Discard the isopropanol, centrifuge the mixture again briefly, and pipette out the remaining isopropanol without disrupting the pellet.

48| Air-dry the pellet at room temperature for 5 min and resuspend the pellet in 250 µl of TE buffer by tapping the microcentrifuge tube, until the DNA is completely dissolved.

49| Add 1 µl of RiboShredder RNase blend to the microcentrifuge tube and incubate at 37 °C for 30 min.

50| Cool the microcentrifuge tube to room temperature and add 250 µl of chilled BACMAX solution 4 to the microcentrifuge tube. Mix by tapping the microcentrifuge tube and incubate on ice for 15 min. Centrifuge the tube at 12,000*g* at 4 °C for 15 min, and transfer the supernatant with a wide-bore tip to a clean microcentrifuge tube without disrupting the pellet.

51| Add 1 ml of absolute ethanol to the supernatant. Mix gently by inverting the microcentrifuge tube a few times. Precipitate the DNA by centrifugation at 12,000*g* at 4 °C for 15 min, and discard the ethanol. Centrifuge the mixture again briefly, and pipette out the remaining ethanol without disrupting the pellet.
▲ **CRITICAL STEP** Because the quantity of the DNA is low, the pellet may not be visible after centrifugation. Keep the tubes in consistent orientations, or label the tubes to mark the location of the pellets.

52| Air-dry the pellet at room temperature for 5 min. Resuspend the pellet by adding 25 µl of TE buffer, tap the microcentrifuge tube a few times, and allow it to remain still at room temperature for 10 min. Use 1 µl to measure the yield of BAC DNA by NanoDrop, which typically should be 30–80 ng μl^{-1}.
■ **PAUSE POINT** The plasmids can be stored at –20 °C for several months.

53| To validate the size of the purified BAC plasmid, linearize the plasmids by combining the following reaction components on ice and incubate at room temperature for 30 min.

Reagent	Volume (µl) for each colony	Total volume (µl) for 5 colonies
10× ME buffer	2	10
ATP solution (10 mM)	2	10
BAC DNA from Step 52 (~30 ng μl^{-1})	5	—
λ-Terminase (2 U μl^{-1})	0.1	0.5
ddH_2O	10.9	54.5
Total	20	75

54| Add 4 µl of 6× DNA loading buffer to the reaction, and load the mixture in 1% Megabase agarose gel (see Equipment Setup section) along with Lambda Ladder PFG marker. Perform PFGE in 0.5× TBE buffer with the CHEF Mapper XA system set to auto-algorithm program with 5–250 kb parameters (6 V cm^{-1}, 0.22–21.79 s, 15 h 16 min, 120°, circulation at 14 °C).

55| After the PFGE, stain the gel with SYBR Gold and visualize the DNA bands using a ChemiDoc XRS+ Imaging System (or equivalent).
? TROUBLESHOOTING

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

Step	Problem	Possible reason	Solution
10	Low sgRNA concentration	Low IVT efficiency or RNase contamination	Check the sequence of the T7 transcription template and re-perform IVT. Pay particular attention to RNase contamination. Denaturing PAGE can be used to assess the degradation of sgRNA
Box 1	No clear band after in-gel Cas9 digestion and PFGE	Low genomic DNA concentration in agarose plugs or poor Cas9 activity	Re-prepare the genomic DNA agarose plugs. Pay particular attention to the starting cell concentration. Use PCR and Sanger sequencing to validate the target sequence in the genome. Check the nuclease activity on short (PCR-obtained) DNA fragments
	Heavy smear after in-gel Cas9 digestion and PFGE	Insufficient cleavage specificity of the RNA-guided Cas9 nuclease (as a result of low sgRNA concentration), too-high cell concentration, or contamination by DNase	Check the concentration and integrity of sgRNA by PAGE, and re-prepare sgRNA if necessary. Pay particular attention to sgRNA concentration (sgRNA concentration >30 ng μl^{-1} is recommended) and RNase contamination. Optimize the conditions by testing a series of different cell concentrations. *Streptomyces* PFGE gels may be more smeary than those of other microorganisms (owing to genomic DNA degradation), in which case the addition of thiourea during PFGE to scavenge reactive Tris radicals may help to reduce the smears. If DNase contamination in the CHEF system is suspected, flush the system thoroughly with bleach and rinse with water
37	Arching during electrotransformation	Too much salt in the electrotransformation buffer or too-high temperature during electroporation	Use less than 2 µl of the ligation mix for electrotransformation. Perform the electroporation in a cold room and prechill the cuvettes on ice
38	No bacteria growth on selective LB agar plates	Poor ligation efficiency, poor electrocompetent *E. coli* cell viability, poor electrotransformation efficiency, incorrect antibiotics used, or too high an antibiotic concentration	Check the vector sequence and pay particular attention to the 30-bp overhang sequences. Use pUC19 control DNA for a control experiment to test the viability of the electrocompetent *E. coli* cells and the efficiency of the electrotransformation experiment. If the time constant of electroporation is shorter than 4 ms, check the device settings and perform the electroporation in a cold room. Double-check the antibiotics and the concentration used on selective LB agar plates
40, 55	No positive clones or very low positive rate	Incorrect ligation or template plasmid contamination	Sequence the plasmids obtained from the colonies grown on selective LB agar plates. Re-design the cloning vector to avoid long identical sequences (≥6 bp) on the vector. If template plasmid contamination is suspected, use the improved method for preparing cloning vectors with long-insert plasmid templates (see Experimental design)

● TIMING

Steps 1–11, day 1, sgRNA template preparation and *in vitro* transcription: 6 h
Steps 12–19, day 1–2, agarose plug preparation and in-gel cell lysis: 2 d
Steps 20–28, day 3, in-gel Cas9 digestion: 3 h
Steps 29–35, day 3, preparation of cloning vectors: 4 h
Steps 36–38, day 3, ligation and electrotransformation: 2 h
Steps 39–55, day 4, validation of positive clones: 1 d
Box 1, day 3, PFGE analysis of cleaved DNA segments (optional): 1 d
Box 2, day 3, preparation of electrocompetent *E. coli* cells (optional): 4 h

ANTICIPATED RESULTS

The described protocol enables the targeted isolation and cloning of long bacterial genomic sequences by the in-gel RNA-guided Cas9 digestion of genomic DNA. By using CATCH, we have been able to efficiently clone bacterial genomic sequences from 30 kb to up to 100 kb. The positive rates (i.e., the percentage of colonies examined successfully containing ligated target inserts; see Steps 39–55) seem to be size dependent, typically between 20% (for 100-kb inserts) and 60% (for 50-kb inserts)[19] (see Fig. 4a in ref. 19). During the PFGE assessment of Cas9 cleavage (**Box 1**), off-target bands may be observed (see lanes 1 and 2 in **Fig. 4a**), but usually they do not affect the cloning results. However, the appearance of smears or the absence of correct bands indicates unsuccessful cleavage of the target DNA segment, which typically results in unsuccessful cloning (see **Table 1** for troubleshooting). We recommend picking ~20 colonies (for short inserts) or all of the colonies (for long inserts, which is important because of the lower cloning efficiency) for PCR validation (Step 40) (typically, there are 50–100 colonies on a plate). During PCR validation, both junctions should be examined, and negative clones may appear (clones 4, 7, 9, 10, 12, 17, and 18 in **Fig. 4d**). In most cases, both junctions are either PCR positive or negative; however, on rare occasions, one junction may appear PCR positive but the other one may not (e.g., clone 18 in **Fig. 4d** should be considered a negative clone). So far, we have successfully cloned several large gene clusters from different bacteria into BAC or p15A vectors, including the 78-kb *pks* gene cluster from *Bacillus subtilis*, the 36-kb *jad* gene cluster from *Streptomyces venezuelae*, the 32-kb *ctc* gene cluster from *Streptomyces aureofaciens* (for cleavage results and positive rates, see ref. 19), and the 83-kb *cda* gene cluster from *Streptomyces coelicolor* (**Fig. 4c**) (positive rate ~5%). *Streptomyces* PFGE gels may appear more smeary than those of other microorganisms (owing to genomic DNA degradation), in which case the addition of thiourea during PFGE to scavenge reactive Tris radicals may help to reduce the smears. Future efforts to improve the methodology and to overcome the limitations may be aided by reducing the mechanical shearing during the procedures, by improving the ligation efficiency of Gibson assembly (especially for long DNA fragments), and by improving the specificity of the Cas9 system[33].

ACKNOWLEDGMENTS We thank C. Lou and Y. Ebenstein for their helpful discussions and co-development of the CATCH concept, and C. Lou and X. Zhao for co-development of the CATCH methodology. This work was supported in part by funding from the National Natural Science Foundation of China (grant nos. 31470532 and 91543102 to T.F.Z.), the Ministry of Science and Technology of China (grant no. 2015CB553402 to T.F.Z.), the Tsinghua University Initiative Scientific Research Program (grant nos. 20151080710 and 20161080133 to T.F.Z.), the Tsinghua University-Peking University Center for Life Sciences (CLS), and the Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases.

AUTHOR CONTRIBUTIONS W.J. performed the experiments. W.J. and T.F.Z. wrote the paper.

COMPETING FINANCIAL INTERESTS The authors declare competing financial interests: details are available in the online version of the paper.

1. Keasling, J.D. Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* **14**, 189–195 (2012).
2. Cobb, R.E. & Zhao, H. Direct cloning of large genomic sequences. *Nat. Biotechnol.* **30**, 405–406 (2012).
3. Gronenberg, L.S., Marcheschi, R.J. & Liao, J.C. Next generation biofuel engineering in prokaryotes. *Curr. Opin. Chem. Biol.* **17**, 462–471 (2013).
4. Pfeifer, B.A., Admiraal, S.J., Gramajo, H., Cane, D.E. & Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**, 1790–1792 (2001).
5. Wong, F.T. & Khosla, C. Combinatorial biosynthesis of polyketides—a perspective. *Curr. Opin. Chem. Biol.* **16**, 117–123 (2012).
6. Zhang, H.B. *et al.* Construction of BIBAC and BAC libraries from a variety of organisms for advanced genomics research. *Nat. Protoc.* **7**, 479–499 (2012).
7. Wang, R.Y., Shi, Z.Y., Chen, J.C. & Chen, G.Q. Cloning large gene clusters from *E. coli* using *in vitro* single-strand overlapping annealing. *ACS Synth. Biol.* **1**, 291–295 2012)).
8. Fu, J. *et al.* Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* **30**, 440–446 (2012).
9. Yamanaka, K. *et al.* Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci. USA* **111**, 1957–1962 (2014).
10. Li, M.Z. & Elledge, S.J. SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol. Biol.* **852**, 51–59 (2012).
11. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
12. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
13. Doudna, J.A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
14. Ran, F.A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
15. Auer, T.O., Duroure, K., Concordet, J.P. & Del Bene, F. CRISPR/Cas9-mediated conversion of eGFP- into Gal4-transgenic lines in zebrafish. *Nat. Protoc.* **9**, 2823–2840 (2014).
16. Karvelis, T., Gasiunas, G. & Siksnys, V. Programmable DNA cleavage *in vitro* by Cas9. *Biochem. Soc. Trans.* **41**, 1401–1406 (2013).
17. Kim, J.M., Kim, D., Kim, S. & Kim, J.S. Genotyping with CRISPR-Cas-derived RNA-guided endonucleases. *Nat. Commun.* **5**, 3157 (2014).
18. Lee, N.C., Larionov, V. & Kouprina, N. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res.* **43**, e55 (2015).
19. Jiang, W. *et al.* Cas9-assisted targeting of chromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* **6**, 8101 (2015).
20. Zetsche, B. *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
21. Ran, F.A. *et al.* *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–191 (2015).
22. Guo, H. *et al.* Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat. Protoc.* **10**, 645–659 (2015).
23. Levy-Sakin, M. & Ebenstein, Y. Beyond sequencing: optical mapping of DNA in the age of nanotechnology and nanoscopy. *Curr. Opin. Biotechnol.* **24**, 690–698 (2013).
24. Gu, H. *et al.* Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **6**, 468–481 (2011).
25. Angeles, V.T. & Reijo Pera, R.A. DNA methylation analysis of human imprinted Loci by bisulfite genomic sequencing. *CSH Protoc.* http://dx.doi.org/10.1101/pdb.prot5046 (2008).

26. Hsu, P.D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
27. Bolukbasi, M.F., Gupta, A. & Wolfe, S.A. Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. *Nat. Methods* **13**, 41–50 (2015).
28. Cho, S.W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **24**, 132–141 (2014).
29. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. & Joung, J.K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
30. Keijser, B.J.F., van Wezel, G.P., Canters, G.W., Kieser, T. & Vijgenboom, E. The ram-dependence of *Streptomyces lividans* differentiation is bypassed by copper. *J. Mol. Microbiol. Biotechnol.* **2**, 565–574 (2000).
31. Peterson, K.R. Preparation of intact yeast artificial chromosome DNA for transgenesis of mice. *Nat. Protoc.* **2**, 3009–3015 (2007).
32. Gibson, D.G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–1220 (2008).
33. Slaymaker, I.M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).