STAR Protocols

Protocol

MegaGate: A toxin-less gateway molecular cloning tool

Christian Kramme,
Alexandru M. Plesa,
Helen H. Wang, ...,
Richie E. Kohman,
Pranam Chatterjee,
George M. Church

pranam.chatterjee@wyss.
harvard.edu

Highlights

MegaGate is a toxin-less Gateway molecular cloning tool

MegaGate shows 99.8% cloning efficiency and high colony numbers compared with Gateway

MegaGate can clone 300+ genes into barcoded destination vectors in a single pool

MegaGate is universally compatible with existing pENTR libraries and ORF cloning

Gateway cloning employs the use of the *ccdb* toxin and has low colony numbers, making it difficult to apply at scale to clone libraries of cDNA vectors. In this protocol, we describe MegaGate, a toxin-less Gateway technology capable of robust cDNA library cloning that is efficient, cheap, and scalable. MegaGate eliminates the *ccdb* toxin used in Gateway recombinase cloning and instead utilizes meganuclease-mediated digestion to eliminate background vectors during cloning and is 99.8% efficient with high colony numbers.

STAR Protocols

Protocol

MegaGate: A toxin-less gateway molecular cloning tool

Christian Kramme,[1,2,3,4] Alexandru M. Plesa,[1,2,3] Helen H. Wang,[1,2] Bennett Wolf,[1,2]
Merrick Pierson Smela,[1,2] Xiaoge Guo,[1,2] Richie E. Kohman,[1,2] Pranam Chatterjee,[1,2,5,*]
and George M. Church[1,2]

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

[2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA

[3]These authors contributed equally

[4]Technical contact

[5]Lead contact

*Correspondence: pranam.chatterjee@wyss.harvard.edu
https://doi.org/10.1016/j.xpro.2021.100907

SUMMARY

Gateway cloning employs the use of the *ccdb* toxin and has low colony numbers, making it difficult to apply at scale to clone libraries of cDNA vectors. In this protocol, we describe MegaGate, a toxin-less Gateway technology capable of robust cDNA library cloning that is efficient, cheap, and scalable. MegaGate eliminates the *ccdb* toxin used in Gateway recombinase cloning and instead utilizes meganuclease-mediated digestion to eliminate background vectors during cloning and is 99.8% efficient with high colony numbers.
For complete details on the use and execution of this protocol, please refer to Kramme et al. (2021).

BEFORE YOU BEGIN

The protocol below describes the process of MegaGate cloning. The "before you begin" section additionally describes obtaining or generating MegaDestination vectors and MegaGate-compatible pENTR vectors. pENTR vectors, known as Entry vectors, are vectors commonly used in Gateway cloning, which contain the to-be-cloned gene flanked by AttL sites. These vectors are used with the "LR" reaction of Gateway enzyme mixes to shuttle the gene into the desired destination vector, which has the corresponding AttR sites. Here we discuss using the classical *att* Gateway sites and two meganucleases, I-SceI and I-CeuI, for the MegaGate reaction.

A general schematic of how MegaGate functions is presented in the Figure 1A. Conceptually, MegaGate functions similarly to Gateway in that the lambda phage recombinase, present in Gateway commercial mixes, shuttles a gene of interest from pENTR to pDestination. In Gateway, this shuttle event replaces the *ccdb* toxin, present on the destination vector, with the gene of interest. Therefore, *ccdb*-sensitive bacteria that receive plasmids with the gene of interest grow, while those that receive unreacted pDestination vector die. In MegaGate, bacteria that receive a plasmid with the gene of interest grow under antibiotic selection. However, if a plasmid does not contain a gene of interest, it retains the meganuclease recognition cassette which is digested by the meganucleases in the MegaGate reaction mix. This linearized plasmid is unable to propagate in bacteria, thus bacteria that receive this digested pDestination die during antibiotic selection. Therefore, MegaGate is able to effectively replace the toxin of Gateway with meganuclease restriction digest.

MegaGate is compatible with traditional Gateway donor libraries. The use of meganucleases, whose recognition sequences are highly uncommon naturally, ensures MegaGate is universally compatible
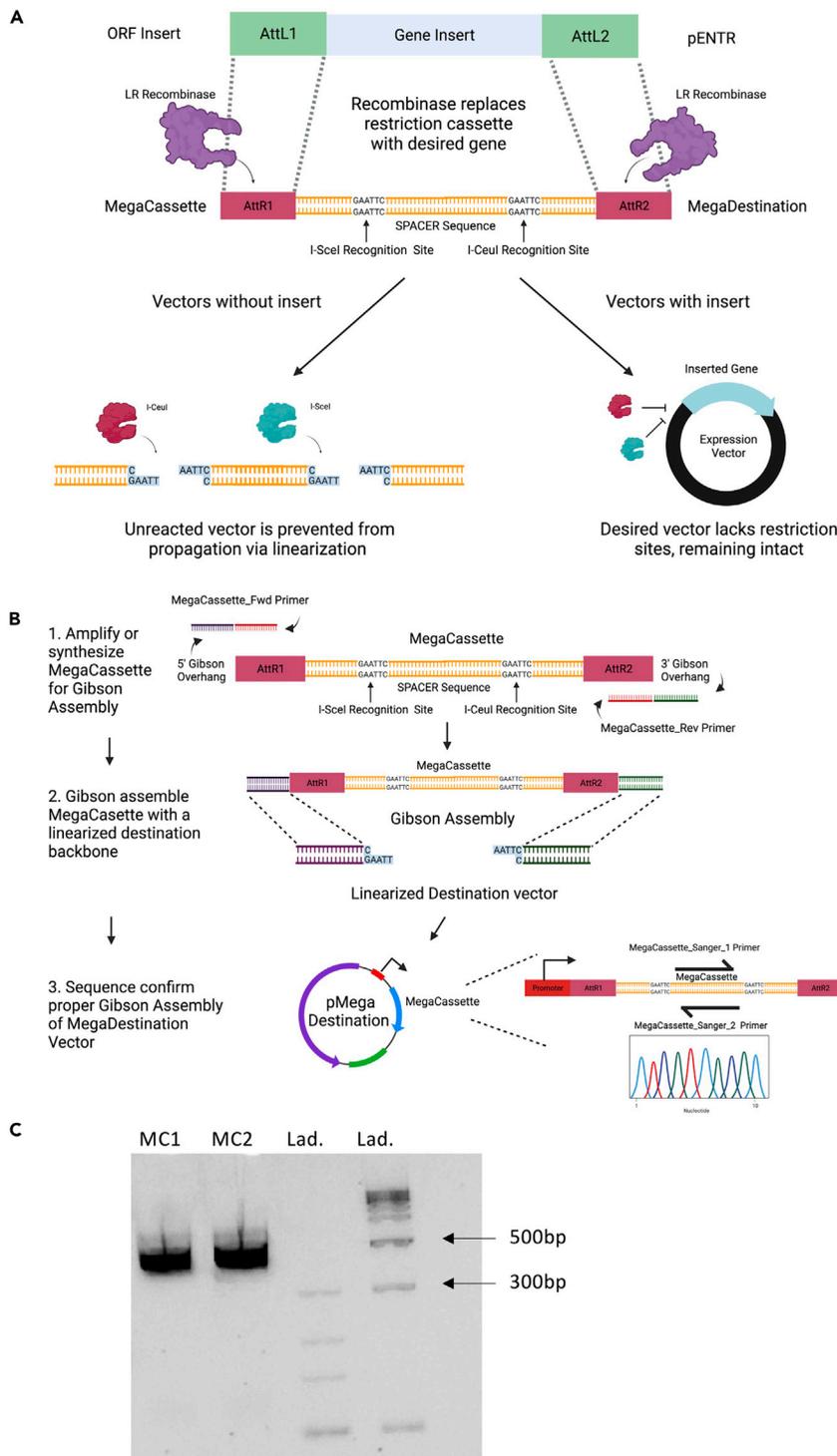
Figure 1. Generating pMegaDestination vectors for cloning
(A) Schematic of MegaGate cloning reaction.
(B) Schematic of steps to generate custom pMegaDestination vectors.
(C) Representative gel image of a MegaCassette PCR reaction. Two Megacassette PCR reactions (MC1 and MC2) were run on a 4% E-Gel for visualization versus the E-Gel Ultra Low range ladder and the E-Gel 1KB Plus DNA Express ladder.

with mammalian open reading frame cloning. In addition, without the toxin, MegaGate vectors can be propagated and easily modified, such as by adding DNA barcodes, in standard bacteria such as NEB 5-Alpha and do not require *ccdb*-resistant bacterial strains that are challenging to clone and work with.

Procure MegaDestination vectors

⏱ Timing: 1–2 days

A variety of inducible and constitutive, piggyBac-compatible MegaDestination vectors are available on Addgene from this publication and are listed in the key resources table (KRT). Additional Mega-Gate-compatible plasmids can be found in our previous publication. (Kramme et al., 2021) These destination vectors are compatible with expression vector cloning for inserting AttL1/AttL2-flanked pENTR-ORFs via MegaGate.

In addition, any expression vector such as 1435 pSG5L Flag HA (Addgene #10791), PB-CA (Addgene #20960), pLEX_305 (Addgene #41390) can be customized into a MegaDestination vector to allow for user-defined screening applications. The MegaCassette, a DNA sequence containing the mega-nuclease restriction sites flanked by AttR1 and AttR2 should be cloned downstream of a desired promoter to create a MegaGate compatible expression vector. See Figure 1B for schematic representation of the below steps for creating custom pMegaDestination vectors. In addition, the full length MegaCassette is provided in the KRT for reference. To generate a custom MegaDestination vector perform the following:

1. Design primers for Gibson Assembly of MegaCassette into custom vector.
   a. Design Gibson overhangs using a tool such as Geneious Gibson Assembly or by creating 20 to 25 bp overhangs on the 5′ and 3′ side of the selected cut sites on the desired expression vector. (Gibson et al., 2009)
   b. Add the 5′ custom overhang to MegaCassette_Fwd primer. Refer to the KRT for primer sequence of MegaCassette_Fwd.
   c. Add the reverse complement of the 3′ custom overhang to the MegaCassette_Rev primer. Refer to KRT for primer sequence MegaCassette_Rev.

   Note: The above primers are for amplification of the MegaCassette, provided in the KRT.

2. Utilize a unique restriction enzyme to linearize the custom expression vector in the position the MegaCassette will be inserted.

   Note: For best results, utilize two different enzymes to generate non-overlapping sticky ends and to remove any unwanted parts of custom vector.

   a. Gel Purify the linearized vector using a 2% agarose gel. Extract the linearized backbone using NEB Monarch Gel Extraction Kit or similar kit.

3. Utilizing a vector containing a MegaCasette, such as Addgene vector #175267, as the input, amplify the MegaCassette with the following reaction conditions.

   Alternatives: MegaCassettes can be synthesized by providers such as IDT. This cassette should be synthesized with the Gibson overhangs present for ease of cloning. If synthesizing, proceed directly to step 6.

Reagent	Amount
Addgene Vector 175267	50 ng
Step 1A Custom MegaCassette fwd Primer (10uM)	1 uL
Step 1B Custom MegaCassette Rev Primer (10uM)	1 uL
2× Q5 Master Mix	12.5 uL
ddH2O	Fill to 25 uL
Total	25 uL

4. Place reaction into thermocycler with the following cycling conditions:

Cycling conditions			
Steps	Temperature	Time	Cycling
Initial Denaturation	98°C	30 s	1
Denaturation	98°C	5 s	25
Anneal	60°C	10 s	
Extension	72°C	10 s	
Final Extension	72°C	2 min	1
Hold	4°C	∞	

5. Gel purify the PCR product on a 2% agarose gel, see Figure 1C for an example gel for this reaction. The correct band can be isolated through gel purification using NEB Monarch Gel purification kit or through bead purification using Promega ProNext beads. Alternative kits for gel purification or bead purification can be utilized.

⚠ CRITICAL: Ensure the product is the correct size (∼400bp for a traditional MegaCassette).

Alternatives: If a large number of off-target bands are present, DMSO can be added to the reaction at 5% to increase specificity.

6. Perform Gibson Assembly.
   a. Combine the gel-purified MegaCassette product with the gel purified linearized backbone vector at a molar ratio of 5:1 insert to backbone.
   b. Add 2× Gibson assembly master mix.
   c. Incubate 15 min to 1 h at 50°C.
   d. Transform into competent *E. coli* such as NEB 5-Alpha and plate on LB agar with antibiotic selection.
   e. Select colonies after 14–16 h 37°C culture and send for Sanger sequencing to confirm proper insertion of cassette.

⚠ CRITICAL: Multiple primers should be used for sequencing confirmation to ensure entire MegaCassette is inserted correctly. AttR sites are challenging to sequence with Sanger, so primers binding inside and outside of the MegaCassette will be necessary for proper sequencing. Primers for sanger sequencing that bind inside the MegaCassette are listed in the KRT as MegaCassette_Sanger 1–2. The binding locations for these primers are depicted in Figure 1B. Additional primers that bind outside of the MegaCassette on the vector backbone will need to be custom generated depending on the application.

Procure pENTR vectors

⊙ Timing: 1–2 days

MegaGate is compatible with traditional Gateway donor libraries known as pENTR libraries. Common examples are the ORFeome and TFome. (Rual et al., 2004; Ng et al., 2020) If available, one may obtain pENTR ORFs by ordering from the ORFeome, TFome or Addgene, specifically ORFs with pENTR or pDONOR in their description.

If a gene is not available, one can simply make a pENTR ORF through the following steps.

7. Design sequence specific primers.
   a. Design a binding site of ∼20 bp on the 5′ and 3′ end of the target sequence.
   b. Add ATTB1 to the 5′ primer and ATTB2 to the 3′ primer, both of which can be found in the KRT.
   c. Amplify the gene of interest using a polymerase such as NEB 2× Q5 Polymerase Master Mix, according to manufacturer's instructions. (https://www.neb.com/protocols/2012/12/07/protocol-for-q5-high-fidelity-2x-master-mix-m0492). Utilize a plasmid containing the gene of interest, genomic DNA or cDNA.

   Alternatives: Synthesize the full-length gene from a provider such as IDT, and place full length ATTB1 and ATTB2 on 5′ and 3′ ends of gene. If synthesizing, use the reverse complement of the ATTB2 sequence provided in the KRT. Proceed directly to step 10.

8. Gel purify the PCR product, using methods described above.
9. Obtain a suitable pDONOR vector such as pDONOR221 from Thermo Fisher Cat #12536017.
10. Perform Gateway BP cloning to insert ORF into pDONOR according to manufacturer instructions. (https://tools.thermofisher.com/content/sfs/manuals/gateway_clonaseii_man.pdf)
11. Transform into competent *E. coli* such as NEB 5-Alpha and plate on LB agar with antibiotic selection and grow 14–16 h at 37°C.
12. Pick colonies and sequence confirm the ORF identity and grow liquid culture to obtain a plasmid preparation of at least 5 ng/ul.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
NEB 5-Alpha	New England Biolabs	Cat: C2987
Chemicals, peptides, and recombinant proteins		
LR Clonase II	Thermo Fisher Scientific	Cat: 11791020
BP Clonase II	Thermo Fisher Scientific	Cat: 11789020
I-SceI (5U/ ul)	New England Biolabs	Cat: R0694
I-CeuI (5U/ ul)	New England Biolabs	Cat: R0699
10× CutSmart Buffer	New England Biolabs	Cat: B7204
T5 Exonuclease	New England Biolabs	Cat: M0663
Gibson Assembly Master Mix	New England Biolabs	Cat: E2611
BsaI-HFV2	New England Biolabs	Cat: R3733
BsmBI-V2	New England Biolabs	Cat: R0739
SapI	New England Biolabs	Cat: R0569
ProNex Size-Selective Purification System	Promega	Cat: NG2001
Critical commercial assays		
Qubit dsDNA HS Assay Kit	Invitrogen	Cat: Q32851
Q5 High Fidelity 2× Mastermix	New England Biolabs	Cat: M0492
Qiagen Plasmid Plus Midi Kit	QIAGEN	Cat: 12941
QIAprep Spin Miniprep Kit	QIAGEN	Cat: 27104
Monarch DNA Gel Extraction Kit	New England Biolabs	Cat: T1020

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw MegaGate pooled cloning NGS data	Kramme et al. (2021)	PRJNA753802
Oligonucleotides		
ORF-BC_Rev: TCTTATCATGTCTGGATCGCGG (For identifying gene-barcode pairs in Figure 3)	This Paper	N/A
MegaCassette_Fwd : ACCACAAGTTTGTACAAAAAAGC	This Paper	N/A
MegaCassette_Rev : ACCACTTTGTACAAGAAAGC	This Paper	N/A
MegaCassette_Sanger_1 : TAAGCGCGCTATGATGGAGG (For sequence confirming the MegaCassette)	This Paper	N/A
MegaCassette_Sanger_2 : ATAGCCTCCATCATAGCGCG (For sequence confirming the MegaCassette)	This Paper	N/A
Recombinant DNA		
pPB-EF1a-MegaGate-DD-Hygro	This Paper	Addgene Plasmid # 175267
pPB-EF1a-MegaGate-HA-Hygro	This Paper	Addgene Plasmid # 175268
pPB-EF1a-MegaGate-DD-Blast	This Paper	Addgene Plasmid # 175270
pPB-EF1a-MegaGate-HA-Blast	This Paper	Addgene Plasmid # 175271
PB-CT3G-ERP2-MG	Kramme et al. (2021)	Addgene Plasmid # 175501
PB-CT3G-cERP2-MG	Kramme et al. (2021)	Addgene Plasmid # 175503
PB-hEF1a-MG-U6-sgRNA	Kramme et al. (2021)	Addgene Plasmid # 175505
PB-CT3G-ERP2-MG-IRES2-mNeonGreen	Kramme et al. (2021)	Addgene Plasmid # 175506
PB-CT3G-CERP2-MG-IRES2-mCherry	Kramme et al. (2021)	Addgene Plasmid # 175507
PB-CT3G-CERP2-MG-IRES2-mNeonGreen	Kramme et al. (2021)	Addgene Plasmid # 175508
PB-CT3G-CERP2-MG-IRES2-mTagBFP2	Kramme et al. (2021)	Addgene Plasmid # 175509
ATTB1 : GGGGACAAGTTTGTACAAAAAAGCAGGCTTA	This Paper	N/A
ATTB2 : GGGGACCACTTTGTACAAGAAAGCTGGGTA	This Paper	N/A
MegaCassette (Sequence is found and annotated in MegaDestination vectors listed under Recombinant DNA. Contains AttR1, I-SceI, I-CeuI, AttR2): ACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAA AATGATATAAATATCAATATATTAAATTAGATTTTGCA TAAAAAACAGACTACATAATACTGTAAAACACAACAT ATCCAGTCACTATGGCGACAGAAGAAGTATAGGGAT AACAGGGTAATTGTTGTAAGCGCGCTATGATGGAGG CTATGCCACTAGAATCTGCGTTCGCTACCTTAGGACC GTTATAGTTAGAAGGAAAGCTCCATCATAGTGACTGG ATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTA TGCAAAATCTAATTTAATATATTGATATTTATATCATTTT ACGTTTCTCGTTCAGCTTTCTTGTACAAAGTGGT	This Paper	N/A
Software and algorithms		
Geneious Prime 2019.2.3	Biomatters Ltd.	N/A
GraphPad Prism v8.3.1 for MacOS	Graph Pad software	N/A

MATERIALS AND EQUIPMENT

Complete MegaGate LR Reaction	
Reagent	Amount
LR Clonase II Enzyme Mix	1 uL
I-SceI (5 U/ul)	1 uL
I-CeuI (5U/ul)	1 uL
10× CutSmart Buffer	5 uL
pENTR vector	50ng
pMegaDestination vector	75ng
ddH2O	Fill to 50 uL
Total	50 uL

STEP-BY-STEP METHOD DETAILS

The conceptual workflow of MegaGate cloning is presented in the Graphical Abstract. The Protocol has three main steps: assemble the reaction, place into thermocycler, transform and sequence colonies.

Assemble the MegaGate reaction

⏱ Timing: 5 min

All components for a MegaGate reaction are combined at once and placed into a thermocycler. The below reaction is for creation of expression vectors by reacting pENTR vectors with pMegaDestination vectors.

1. Assemble the following MegaGate reaction on ice.

Complete MegaGate LR Reaction	
Reagent	Amount
LR Clonase II Enzyme Mix	1 uL
I-SceI (5U/ul)	1 uL
I-CeuI (5U/ul)	1 uL
10× CutSmart Buffer	5 uL
pENTR vector	50ng
pMegaDestination vector	75ng
ddH2O	Fill to 50 uL
Total	50 uL

Alternatives: The same reaction can be used for MegaGate BP reactions, replacing LR Clonase II with BP Clonase II and using the appropriate AttB flanked gene insert and a pMegaDonor vector.

⚠ CRITICAL: If scaling MegaGate reaction down to 25 uL total, it is critical that the reagents are used in the same concentration. If using alternative endonucleases in customized vectors, the amounts and protocols will need to be optimized for the specific nuclease. A single meganuclease can additionally be used instead of two (use 2 uL instead of 1 ulL) with a resulting decrease in efficiency, as is seen in Figure 3D. Additionally, non-meganuclease enzymes such as Golden Gate Type IIS enzymes can be used effectively, as seen in Figure 3G.

Note: If using MegaGate for pooled cloning, total pENTR vector should be 50 ng. pENTR vectors should be pooled in an equimolar fashion to minimize cloning bias (Figures 2C–2E). For best results, keep the deviation in insert size at less than 25% and be aware that smaller inserts clone more efficiently (Figures 2C–2E). Greater than 50 ng of pENTR can be utilized if needed. pENTR vector amounts of less than 10 ng will result in substantial reduction in colonies obtained. For ORFs larger than 5 KB, greater than 50 ng can be utilized to increase efficiency.

Note: pMegaDestination vectors can be pooled, such as barcoded destination vectors (Figures 2C–2E). Pooled destination vectors should not exceed 75ng total to minimize background.

Thermocycle the MegaGate reaction

⏱ Timing: 2 h and 20 min

This step performs the MegaGate reaction entirely within a thermocycler.
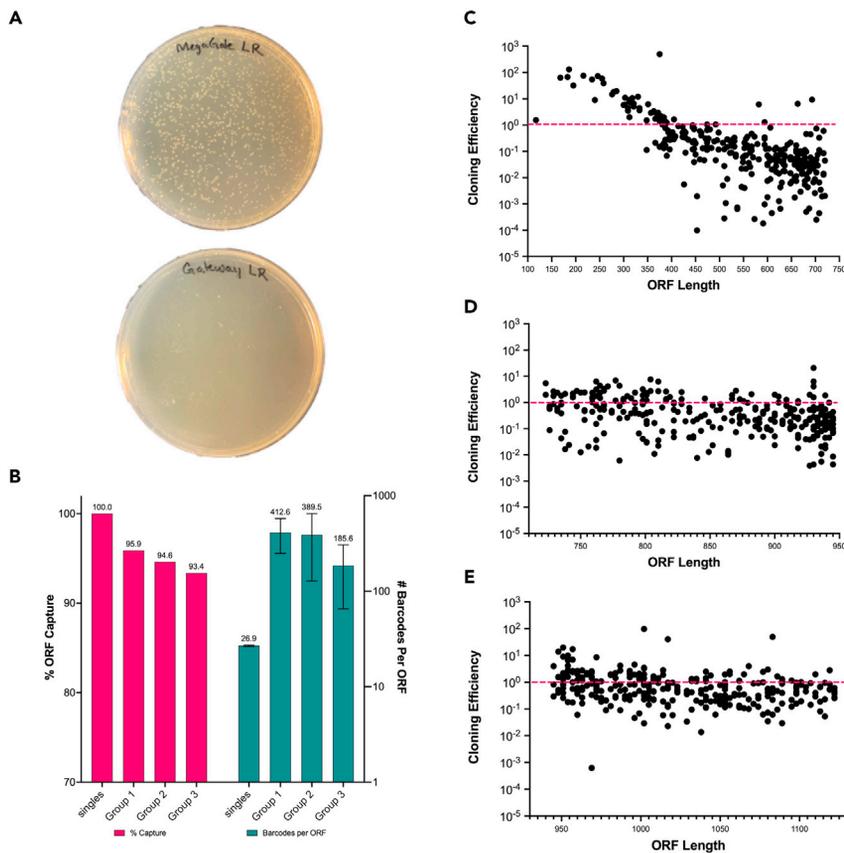
Figure 2. Utilization of MegaGate for single and pooled cloning

(A) Depiction of representative MegaGate transformant plates versus Gateway transformant plate. A pENTR-sfGFP was used for both reactions at 50 ng. Gateway was performed according to manufacturer's instruction using 4 ul of LR Clonase II enzyme mix and TE-8. Gateway and MegaGate were performed using 75 ng of each destination vector, which are equivalent sizes. The reactions were additionally performed utilizing the same aliquot of LR Clonase and NEB 5-alpha cells to ensure there was no bias.

(B) Percent ORF capture (pink) measured as the ratio of input genes captured in expression vectors in a single MegaGate cloning reaction for single genes and pooled groups. Number of barcodes captured per gene (teal) for the single genes and pools was determined via NGS alignment of destination vector amplicons.

(C–E) Cloning efficiency as a function of ORF length. Cloning efficiency is measured as the relative abundance of the gene in the expression vector pool divided by the relative abundance of the gene in the pENTR pool, as measured by NGS counts. Genes are arrayed by length on the x-axis. Each pool contained 300 pENTR genes, pooled by size. 50 ng of total pENTR gene were used for a single MegaGate reaction for each pool. Colonies were transformed and plated on four plates per pool and colonies were scrapped and grown overnight to obtain expression plasmids. PCR of the gene insert on the pENTR and pExpression plasmids was utilized to prepare libraries for NGS, and indexed PCR libraries were run on an Illumina MiSeq. Geneious RNA aligner was used to align reads to ORFs for each pENTR and corresponding pExpression pool and BBMap was used to call barcodes for each ORF read in the expression pool.

2. Insert the assembled MegaGate reaction into thermocycler with the following settings.

Cycling conditions		
Steps	Temperature	Time
Gateway Insertion	25°C	1 h
Restriction Digest	37°C	1 h
Inactivation	65°C	20 min
Hold	4°C	∞

Note: If using custom nucleases not listed in KRT, the manufacturer recommended cutting temperature should be utilized instead of 37°C. In addition, isothermal reactions at 25°C and 37°C can be utilized with a corresponding lowering in overall efficiency (Figure 3E).

⫿⫿ **Pause point:** Reaction can be stored at −20°C for later bacterial transformation or proceed directly to transformation. The reaction is stable for up to one week; however, more testing is needed to determine the stability of the reaction for greater than one week.

Transform and sequence the MegaGate reaction

⏲ Timing: 2 days

This step performs the transformation of the reaction into NEB 5-Alpha bacteria, with subsequent incubation and colony picking.

3. Transform 2 uL of the finished MegaGate reaction into 10 uL of NEB 5-Alpha cells according to manufacturer's instructions. https://www.neb.com/protocols/0001/01/01/high-efficiency-transformation-protocol-c2987.

   Note: Any standard bacterial strain can be used for cloning and propagation.

4. Plate reaction onto LB agar plate with proper antibiotic and grow 14–16 h at 37°C.
5. Pick single colonies for individual cloning reactions or alternatively scrape plates for library preparations.
6. Grow colonies in liquid broth (LB) media with antibiotic 12–16 h at 37°C to propagate plasmids.
7. Harvest plasmids from bacteria using plasmid preparation kits such as QIAPrep Spin MiniSpin kit.
8. Sequence plasmids using Sanger sequencing or NGS to confirm insertion. If using barcoded versions of MegaDestinations from Kramme et al. (2021) or this paper, one can use the ORF-BC_Rev primer in the KRT for Sanger sequencing to identify the barcode and inserted gene simultaneously.

EXPECTED OUTCOMES

A typical MegaGate transformant plate versus a Gateway transformant plate is shown in Figure 2A. As can be seen, MegaGate often obtains far higher colony numbers compared to Gateway. Expected outcomes from single gene and pooled (300+ ORF) cloning to barcoded destination vectors are shown in Figures 2B–2E. Figure 2B provides summary statistics for single and pooled cloning reactions, represented as percent ORF capture, which is measured as the ratio of pENTR ORFs successfully cloned into destination vectors, and barcodes per ORF, which is measured via Sanger sequencing or next generation sequencing (NGS) of unique barcodes per ORF in the expression vector pool. Figure 2B is further expanded on in Figures 2C–2E, which show the cloning efficiency for the three separate pooled reactions of 300 ORFs per pool, which were grouped by size. As can be seen in Figure 2C, when the standard deviation of ORF sizes in the pool is large, larger ORFs clone less efficiently than smaller ORFs, seen as a decrease in efficiency as ORF size increases for this pool. When ORF size deviation is reduced, as is the case for pool 2 and 3 in Figures 2D and 2E, there is a flat distribution of cloning efficiency in which the relative ratio of the ORFs in the pENTR pool is similar to the relative ratio of ORFs in the resulting expression pool.

Additionally, cloning efficiency outcomes from alternate reaction conditions are shown in Figure 3. In order to optimize the MegaGate reaction, we systematically tested cutting duration (Figure 3A), reaction volume (Figure 3B), unique reaction steps (Figure 3C), units of cutting enzyme in reaction (Figure 3D), reaction temperature (Figure 3E), number of restriction digest sites on vector (Figure 3F), restriction enzyme type (Figure 3G), and the use of BP Clonase in MegaGate versus Gateway
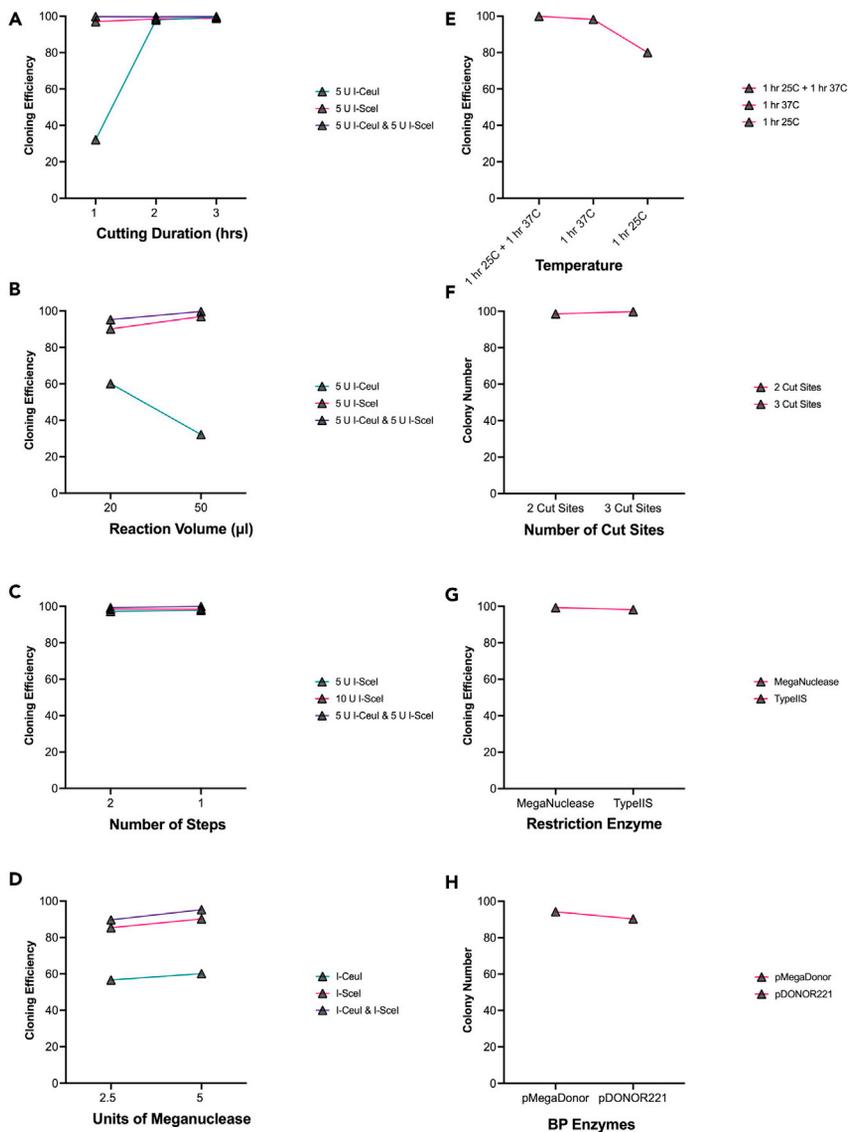
Figure 3. Cloning efficiency optimization of MegaGate cloning reactions

(A and B) Cloning efficiency is determined as the number of colonies on the plus insert plate compared to the total number of colonies on the minus insert plate. Reaction conditions tested are as follows: (A) endonuclease cutting time (B) total reaction volume.

(C) All components added at once (one step) reaction versus separate recombinase and endonuclease reactions (two step).

(D) Units of meganuclease added.

(E) Multi-temperature reactions versus isothermal.

(F) Number of restriction sites on MegaDestination vector.

(G) Meganuclease versus Type IIS endonuclease.

(H) Gateway BP reaction versus MegaGate BP reaction colonies obtained.

(Figure 3H). We additionally tested these reaction conditions for their total obtained colony numbers, which can be seen in Figure 4. Based on these optimization reactions, we determined that an all-in-one reaction mix with a three-step temperature change in a 50 uL reaction was most optimal for Megagate cloning in terms of cloning efficiency and colonies obtained, which is described above. Our results also show the MegaGate reaction can be effectively customized, such as changing recombinase or restriction endonucleases, to fit specific cloning needs. Refer to
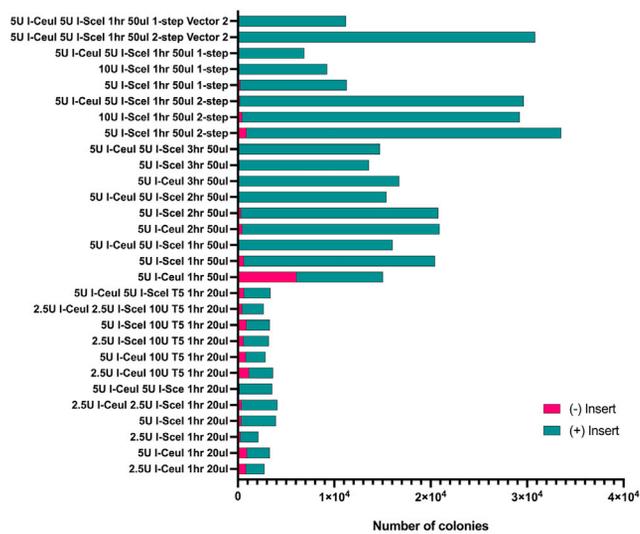
Figure 4. Colony number optimization of MegaGate cloning reactions
Colony numbers obtained from different MegaGate optimization cloning reactions are shown with colonies obtained on the x-axis and the condition on the y-axis. Colony number was determined by counting colonies from a 2 uL transformation and multiplying by 25 to account for the total colonies in the 50 uL MegaGate reaction. Colony numbers from the plus insert condition are shown in teal and colonies from the minus insert condition (background colonies) are shown in pink.

Kramme et al. (2021) for further applications of MegaGate cloning within the STAMPScreen mammalian genetic screening pipeline.

The results of a single MegaGate cloning reaction are generally 99.8% cloning efficiency, measured as the relative ratio of colonies containing the desired insert versus colonies without the insert. The total colony number varies based on destination vector and gene insert. On average, inserts of around 1 kb will yield 100–300 colonies from a single 2 uL transformation, as can be seen in Figures 2A and 4. Smaller inserts will yield 1000+ colonies and larger inserts (10 kb +) will yield around 10–50 colonies. Background colonies i.e., those lacking an insert, will appear at roughly 0.2% on average when the reaction is performed optimally. Modifications of the protocol such as scaling the reaction down, limiting insert amount, excessive destination amount and altered cycling times or temperatures will result in higher background or lower overall colony numbers. When MegaGate is used for pooled cloning of 300 + ORF inserts, we expect a roughly 95% ORF capture on the first attempt if the pENTR library is maintained as roughly equimolar (Figure 2B). Large differences in pENTR-ORF frequency can result in loss of lower abundance ORFs. In addition, for pooled cloning, ORFs should be pooled based on size. Libraries with large differences in size can enrich smaller ORFs which clone more efficiently than larger ORFs (Figures 2C–2E). For MegaGate reactions using the BP Clonase system, expect on average 5–50 colonies per ORF. This reaction is less efficient overall than the LR reaction and may require further optimization.

MegaGate is an enabling technology for use in cDNA screening and cell engineering for mammalian systems. We previously detailed an all-in-one pipeline for mammalian genetic screening known as STAMPScreen (Kramme et al., 2021), in which MegaGate is utilized to generate cDNA screening libraries. MegaGate improves on traditional Gateway cloning by elimination of the *ccdb* cassette while maintaining compatibility with ORF libraries such as the ORFeome and does not suffer from the ORF PCR amplification or restriction site compatibility constraints of Golden Gate and Gibson cloning (Magnani et al., 2006; Engler et al., 2008; Gibson et al., 2009). MegaDestination vectors can optionally feature unique DNA barcodes that can be captured through gDNA barcode sequencing (BAR-Seq), as well as captured alongside the transcriptome through RNA-Seq,

scRNA-Seq, and targeted RNA-Seq which are common in many NGS-based screening pipelines (Smith et al., 2009; Martin et al., 2016). We show that MegaGate cloning is efficient at cloning more than 1000 human genes individually or in pools and linking each gene with a unique barcode that can be determined through NGS or Sanger sequencing. The expression vectors are then easily integrated into the genomes of hiPSCs or primary fibroblasts to generate stable cell lines with doxycycline inducible genetic cassettes (Yusa et al., 2011). We additionally show in STAMPScreen the development of dual expression systems in which a sgRNA can be cloned into the MegaDestination vector using Golden Gate and a subsequent cDNA is inserted via MegaGate. This technology can be utilized to repress or induce native genes when paired with CRISPRi and CRISPRa tools while simultaneously overexpressing the cDNA construct (Yeo et al., 2018; Chavez et al., 2016). We envision this form of MegaGate cloning to effectively pair genome-wide sgRNA libraries with genome-wide cDNA libraries for library-by-library dual screening. When paired with the computational tools for target prediction and transcriptomic analysis, MegaGate is an effective tool for generating expression libraries for high-throughput screening.

LIMITATIONS

MegaGate is generally universal in its application with only a few limitations. For cloning purposes, genetic sequences that contain the selected meganuclease restriction sites cannot be cloned with MegaGate; however, these restriction site sequences are extremely rare and almost never naturally both present in open reading frames. When one meganuclease cannot be used for this reason, one may utilize twice the amount of the other available meganuclease to obtain positive colonies, albeit with lower efficiency. In addition, we have shown that other restriction endonucleases such as Type IIS enzymes function equally to the meganucleases and can be used as substitutes if a meganuclease cannot be used (Figure 3G). In addition, MegaGate has been extensively tested and optimized with nucleases from New England Biolabs. We have not optimized or tested the use of isoforms and buffers of these nucleases from other providers such as Thermo Fisher. We do not anticipate there being issues and fully expect the technology to be compatible with other enzymes and enzyme providers, but further research is needed.

TROUBLESHOOTING

Problem 1

Low or no colonies obtained from MegaGate LR cloning

A MegaGate reaction may show very low or no colonies after transformation step (Step 3).

Potential solution

If no colonies are obtained, ensure that the selection plate has the appropriate concentration of antibiotic by plating of a positive control. Additionally, when using LR Clonase II and BP Clonase II it is crucial these reagents are always kept on ice as they are extremely sensitive to degraded performance. The most common cause of MegaGate failure is a poorly performing batch of LR Clonase II or BP Clonase II. Ensure the batch is not degraded via a positive control provided by the manufacturer and if it is, obtain a new batch. Additionally, sequence confirm the AttR sites of the MegaDestination and the AttL sites of the pENTR vector to ensure there are no mutations preventing recombinase function. Some ORFs, particularly those greater than 5 kb benefit from increased amounts of pENTR input, therefore increasing the input to 75 ng or more may increase efficiency. One may also transform a larger amount of final product to obtain colonies from poorly performing reactions. Although rare, ensure the insert sequence and backbone of any customized MegaDestination vector do not contain I-SceI or I-CeuI meganuclease restriction sites. In the event all solutions fail, one may try to split the reaction into two parts. First add all MegaGate components except for CutSmart and the meganucleases and place to thermocycler for 1 h at 25°C. Then add CutSmart and the meganucleases and incubate at 37°C for 1 h then 65°C for 20 min. Splitting the reaction to two parts has been shown to generally increase colony number and shows equivalent cloning efficiency (Figure 2C).

Problem 2

Large number of background colonies obtained from MegaGate LR cloning

If a negative control is used, where no pENTR is provided, on average there will be 1 to 10 colonies on the plate at a roughly 200-fold reduction compared to the positive control. Increased numbers of background colonies seen in the sequencing step (Step 8) are a sign of an issue with the reaction.

Potential solution

Ensure that the antibiotic selection plates utilized have a sufficiently high concentration of antibiotic and are not degraded. For the provided MegaDestination vector, we utilize carbenicillin at 100ug/mL and for pMegaDONOR221 we utilize kanamycin at 50ug/mL. Additionally, ensure that both meganucleases were added to the reaction and that the mixture is thoroughly vortexed. Meganucleases are sensitive to degradation, particularly I-SceI, which is stored at −80°C. If enzymes have been subjected to more than 40 freeze thaws, consider using fresh batches to ensure the enzymes are fully functional. One should also sequence confirm the integrity of the MegaCassette in the MegaDestination to ensure that both restriction sites are free of mutations. If large background still exists, increase the amount of meganuclease in reaction and extend the cutting time (37°C step) to 2 h.

Problem 3

Meganuclease recognition site is present in my ORF of interest

If a meganuclease (I-SceI or I-CeuI) recognition site is present in the ORF of interest, the MegaGate reaction will fail to clone, due to the expression vector being linearized by the meganucleases, resulting in no colonies after the transformation step (Step 3).

Potential solution

If only one of two meganuclease recognition sites is present in the ORF, do not include that meganuclease in the MegaGate reaction. Increase the remaining meganuclease to 2 ul. The reaction will likely be overall less efficient, particularly if only I-CeuI is utilized, but will still yield positive colonies. If both meganucleases have recognition sites in the ORF, one will need to clone a new MegaCassette into the expression vector, as described in above in the before you begin section. Alternate endonucleases such as Type IIS Golden Gate enzymes function equivalently to meganucleases and can be utilized as replacements. The meganuclease recognition site can be removed from the gene of interest through codon optimization or site directed mutagenesis.

Problem 4

MegaGate BP Reaction shows little to no transformant colonies

The MegaGate BP reaction in general shows 10–100 fold less colonies obtained compared to the MegaGate LR reaction. More research is needed to determine the cause of this low efficiency. After performing a MegaGate BP reaction, one may obtain no colonies after transformation step (Step 3).

Potential solution

Since BP reactions are commonly used to generate pENTR vectors from a single pDONOR vector (not pools of unique pDONOR vectors), one can simply use the Gateway BP reaction with pDONOR221 or pDONOR223 to generate the vector using the manufacturer's instructions. If using MegaGate BP reaction, one can increase the amount of input PCR product to greater than 100 ng in order to obtain more positive colonies. Ensure that the BP Clonase being utilized is fresh, with limited freeze-thaw cycles. Additionally, transform more of the overall reaction to NEB 5-Alpha bacteria to obtain more transformants.

Problem 5
Sanger Sequencing of MegaGate gene inserts in expression vectors show "no priming", "early termination" or "high background"

Sanger sequencing (Step 8) of picked colonies may show issues. When sanger sequencing of expression vectors shows one of the above issues, it is likely that the colony sequenced is either a background no-insert colony, a primer-dimer insert, or a colony with multiple plasmids.

Potential solution
For early termination errors, the colony is most likely to be a background colony containing no insert. These colonies often show early termination, due to the sanger sequencing prematurely terminating when sequencing across two full length ATT sites, which are repetitive in nature and contain secondary structures. If a large number of early termination colonies is obtained, there is likely an issue with the MegaGate reaction set up. Refer to problem 2 troubleshooting to lower background colonies. If the no priming error occurs, the issue is likely with the sequencing facility or a MegaGate BP reaction. In our experience, the "no priming" error often results from insufficient colony picking during sequencing or from the use of poorly performing sequencing primers. Submission of plasmid preparations (versus colonies) has been shown to fix this issue. For MegaGate BP reactions or Gateway BP reactions, excessive "no priming" results is often caused by primer-dimers being cloned into the pENTR vectors. These primer-dimers, which arise from the PCR amplification of the ORF of interest, must be removed by gel extraction. Failure to remove these will lead to the dimers being cloned into the pDONOR and occurs at a high frequency due to the small size of the dimers, which clone more efficiently. To correct the issue, gel extract the band of interest prior to cloning. For the high background error, ensure that the colonies are not plated densely such that colonies are merging or multiple colonies are picked at once. If error continues, lower the amount of reaction transformed from 2ul to 0.2ul to ensure multiple plasmids are not present in individual bacteria.

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Pranam Chatterjee, pranam.chatterjee@wyss.harvard.edu.

Materials availability
Plasmids generated in this study and referenced in Kramme et al. (2021) have been deposited to Addgene. Refer to KRT for specific catalog numbers.

Data and code availability
Additional data and code can be found in Kramme et al. (2021).

DECLARATION OF INTERESTS

C.K., A.M.P., and G.M.C. are listed as inventors on US Provisional Patent Application 63/148,656 entitled "Compositions and Methods for Molecular Cloning."

REFERENCES

Chavez, A., Tuttle, M., Pruitt, B., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S., Cecchi, R., Kowal, E., Buchthal, J., et al. (2016). Comparison of cas9 activators in multiple species. Nat. Methods 13, 563–567.

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. PLoS One 3, e3647.

Gibson, D.G., Young, L., Chuang, R., Venter, J., Hutchinson, C., and Smith, H. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345.

Kramme, C., Plesa, A., Wang, H., Wolf, B., Smela, M., Guo, X., Kohman, R., Chatterjee, P., and Church, G. (2021). An integrated pipeline for mammalian genetic screening. Cell Rep. Methods 1, 100082.

Magnani, E., Bartling, L., and Hake, S. (2006). From gateway to multisite gateway in one recombination event. BMC Mol. Biol. 7, 46.

Martin, D.P., Miya, J., Reeser, J.W., and Roychowdhury, S. (2016). Targeted RNA sequencing assay to characterize gene expression and genomic alterations. J. Vis. Exp. 114, 54090.

Ng, A.H.M., Khoshakhlagh, P., Arias, J., Pasquini, G., Wang, K., Swiersy, A., Shipman, S., Appleton, E., Kiaee, K., Kohman, R., et al. (2020). A comprehensive library of human transcription factors for cell fate engineering. Nat. Biotechnol. 39, 510–519.

Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P., et al. (2004). Human ORFeome version 1.1: a platform for reverse proteomics. Genome Res. 14, 2128–2135.

Smith, A.M., Heisler, L., Mellor, J., Kaper, F., Thompson, M., Chee, M., Roth, F., Giaever, G., and Nislow, C. (2009). Quantitative phenotyping via deep barcode sequencing. Genome Res. 19, 1836–1842.

Yeo, N.C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C., Guo, X., Sharma, S., Tung, A., et al. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat. Methods 15, 611–616.

Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive piggyBac transposase for mammalian applications. Proc. Natl. Acad. Sci. U S A 108, 1531–1536.