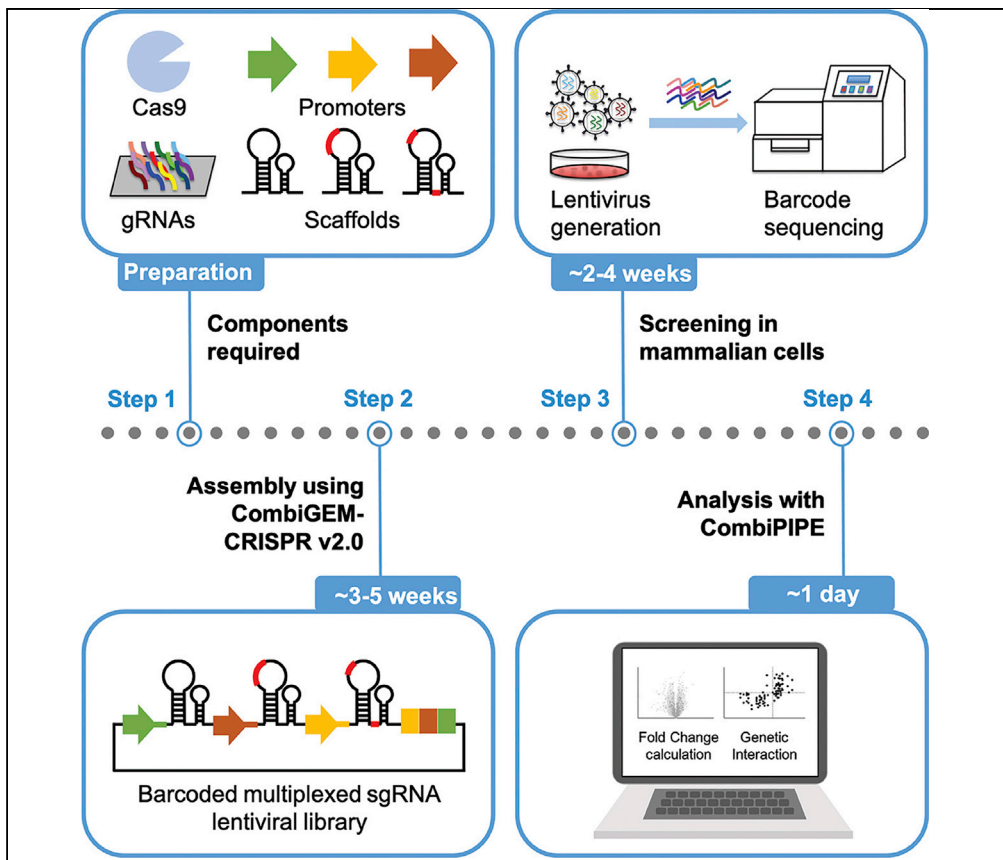


## Protocol

# Extensible combinatorial CRISPR screening in mammalian cells



The CRISPR-Cas system coupled with Combinatorial Genetics *En Masse* (CombiGEM) enables systematic analysis of high-order genetic perturbations that are important for understanding biological processes and discovering therapeutic target combinations. Here, we present detailed steps and technical considerations for building multiplexed guide RNA libraries and carrying out a combinatorial CRISPR screen in mammalian cells. We also present an analytical pipeline, CombiPIPE, for mapping two- and three-way genetic interactions.

Peng Zhou, Yuk Kei Wan, Becky K.C. Chan, Gigi C.G. Choi, Alan S.L. Wong

aslw@hku.hk

### HIGHLIGHTS

Engineered promoter and scaffold sequences for multiplexed guide RNA expression

Extensible assembly and screening of barcoded multiplexed guide RNA libraries

CombiPIPE established for analyzing CombiGEM-CRISPR screen results

Zhou et al., STAR Protocols 2, 100255  
March 19, 2021 © 2020 The Author(s).  
<https://doi.org/10.1016/j.xpro.2020.100255>



## Protocol

## Extensible combinatorial CRISPR screening in mammalian cells

Peng Zhou,<sup>1</sup> Yuk Kei Wan,<sup>1</sup> Becky K.C. Chan,<sup>1</sup> Gigi C.G. Choi,<sup>1</sup> and Alan S.L. Wong<sup>1,2,3,4,\*</sup><sup>1</sup>Laboratory of Combinatorial Genetics and Synthetic Biology, School of Biomedical Sciences, The University of Hong Kong, Pokfulam, Hong Kong SAR, China<sup>2</sup>Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China<sup>3</sup>Technical contact<sup>4</sup>Lead contact\*Correspondence: [aslw@hku.hk](mailto:aslw@hku.hk)<https://doi.org/10.1016/j.xpro.2020.100255>

## SUMMARY

The CRISPR-Cas system coupled with Combinatorial Genetics *En Masse* (CombiGEM) enables systematic analysis of high-order genetic perturbations that are important for understanding biological processes and discovering therapeutic target combinations. Here, we present detailed steps and technical considerations for building multiplexed guide RNA libraries and carrying out a combinatorial CRISPR screen in mammalian cells. We also present an analytical pipeline, CombiPIPE, for mapping two- and three-way genetic interactions.

For complete details on the use and execution of this protocol, please refer to Zhou et al. (2020).

## BEFORE YOU BEGIN

## Overview

The CombiGEM-CRISPR v2.0 toolkit has storage vectors harboring hH1/hU6/mU6 with modified ends and other engineered parts including wild-type/v1/v2 sgRNA scaffold sequences. These part variants are used to reduce sequence similarity among multiple sgRNA expression units to minimize vector recombination. The parts are designed in a way that only one set of barcoded sgRNA oligo pairs is required to generate different sgRNA expression units carrying different promoters and scaffold variants. Here, the protocol starts with how sgRNA are designed and annealed oligos are mixed in a pool for a ligation of the one-way sgRNA storage libraries. Then, we set out the steps for the successive one-pot insertion of each one-way sgRNA expression library into the destination vectors to give a higher-order combinations of sgRNAs. Lastly, we describe how the library is delivered to mammalian cells for multiplexed CRISPR screens, and the subsequent steps for sample preparation for barcode sequencing and data analysis using CombiPIPE.

## Design sgRNA spacer-containing oligos for library construction

Multiple sgRNA design tools have been developed (Chuai et al., 2017). Here we use the GPP sgRNA Designer (Doench et al., 2016; Sanson et al., 2018) as an example to illustrate sgRNA selection.

1. Submit job for designing sgRNAs for the genes of interest following the instructions on the web site: <https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design>.

△CRITICAL: The parameter "Pick Quota" could be left unchecked while we recommend checking the box "Report Unpicked Sequences?" for further selection of sgRNAs based on the criteria listed in step 2.



2. When the job is completed, two result files are generated. “sgRNA Picking Results” lists out all candidate sgRNAs designed by the software, where “sgRNA Picking Summary” lists out the selected sgRNAs according to the “Pick Quota.” For CombiGEM-CRISPR v2.0 experiments, we recommend further selection of sgRNAs from the lists based on the following criteria:
  - a. With on-target efficacy score > 0.6 based on Azimuth 2.0.
  - b. Exclude sgRNAs that are lying too close to target sites of other selected ones (Spacing Violation).
  - c. Exclude sgRNAs containing restriction digestion sites for the plasmid library construction (e.g., BamHI, EcoRI, BglII, MfeI, and BpI used in this cloning scheme).

**Note:** We observed that the on-target efficacy score largely predicts the efficiency of sgRNAs in OVCAR8-ADR cells (Zhou et al., 2020). A score of > 0.6 is a good start. Based on the above criteria, lists of pre-selected sgRNAs for all human and mouse coding genes are included in Tables S1 and S2. Using wild-type SpCas9, sgRNAs with off-target ranks < 100 can be selected. When a high-fidelity SpCas9 variant (such as Opti-SpCas9 (Choi et al., 2019)) that is compatible with sgRNAs containing a starting 5' guanine preferred for transcription initiation under U6 promoter is employed, sgRNAs with higher on-target scores could be selected despite they may have higher off-target ranks. sgRNAs with multiple perfect or near-perfect target sites can also be identified using GuideScan (Perez et al., 2017; Perez et al., 2019) and discarded. sgRNAs can be further selected based on whether they target the sites locating within 5%–65% of the protein-coding sequence or functional domains of the protein-coding sequence (Shi et al., 2015), and the predictions to generate high percentage of frameshift mutations (Allen et al., 2018; Shen et al., 2018). The information relating to criteria b. and “sites locating within 5%–65% of the protein-coding sequence” can be found in column “Picking Notes” of the sgRNA-designs table (output of GPP sgRNA Designer).

3. Assign 8-nucleotide (nt) randomized barcode sequences for each sgRNA.

**△ CRITICAL:** We recommend the GC content of the barcodes to be around 50% and allow at least 2-nt differences to differentiate true identities of sgRNAs from PCR or sequencing errors. A list of compatible barcodes that was used in our previous study (Zhou et al., 2020) is included in Table S3.

4. Design oligo sequences for synthesis (oligos synthesized in this study are DSL-purified). The forward oligo sequence is 5'-CACCG-N(20)-GTTGGGTCTTCGAGAAGACCTATTC-N(8)-C-3'; the reverse oligo sequence is 5'-AATTG-N(8)-GAATAGGTCTTCTCGAAGACCCAAAC-N(20)-C-3', where N(20) represents the sgRNA spacer and N(8) is barcode sequence. The linker between sgRNA and barcode contains two BpI digestion sites (Figure 1).

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
T4 DNA ligase	New England Biolabs	Cat#M0202T
T4 DNA ligase reaction buffer	New England Biolabs	Cat#B0202S
T4 polynucleotide kinase (T4 PNK)	New England Biolabs	Cat#M0201S
FastDigest BpI (BbsI)	Thermo Scientific	Cat#FD1014
FastDigest MnlI (MfeI)	Thermo Scientific	Cat#FD0753
FastDigest green buffer	Thermo Scientific	Cat#B72
FastDigest BglII	Thermo Scientific	Cat#FD0083
FastDigest BamHI	Thermo Scientific	Cat#FD0054
FastDigest EcoRI	Thermo Scientific	Cat#FD0274

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SYBR Safe DNA Gel Stain	Invitrogen	Cat#S33102
UltraPure H <sub>2</sub> O	Invitrogen	Cat#10977015
DMEM	Gibco	Cat#10569010
Fetal bovine serum (FBS)	Gibco	Cat#10270106
Antibiotic-antimycotic (100×)	Gibco	Cat#15240062
Opti-MEM	Gibco	Cat#31985062
FuGene HD transfection reagent	Promega	Cat#E2312
<b>Bacterial strains</b>		
XL10-Gold ultracompetent cells	Agilent	Cat#200317
<b>Critical commercial assays</b>		
Phusion High-Fidelity PCR Master Mix	New England Biolabs	Cat#M0531S
Quant-iT PicoGreen dsDNA Assay kit	Life Technologies	Cat#P11496
Kapa HiFi Hotstart Ready-mix	Kapa Biosystems	Cat#KK2602
5× KAPA HiFi Fidelity w/ Mg <sup>2+</sup> buffer	Kapa Biosystems	Cat#KB2500
Agencourt AMPure XP beads	Beckman Coulter	Cat#A63881
Kapa SYBR Fast qPCR Master Mix	Kapa Biosystems	Cat#KK4600
NEBNext Library Quant Kit for Illumina	New England Biolabs	Cat#E7630S
QIAGEN DNeasy Blood and Tissue kit	QIAGEN	Cat#69506
QIAGEN Plasmid Midi kit	QIAGEN	Cat#12945
QIAGEN Plasmid Maxi kit	QIAGEN	Cat#12162
<b>Oligonucleotides</b>		
WT sgRNA scaffold F: 5'-GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTGGATCCGCAACGGA-3'	This study	N/A
WT sgRNA scaffold R: 5'-GAATCCGTTGCGGATCCAAAAAGCAC CGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTA TTTAACTTGCTATTTCTAGCTCTA-3'	This study	N/A
v1 sgRNA scaffold F: 5'-GTTTTAGAGGATGAAAATCCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTGGATCCGCAACGGA-3'	This study	N/A
v1 sgRNA scaffold R: 5'-GAATCCGTTGCGGATCCAAAAAGCAC CGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTA TTTAACTTGATTTTCATCCTCTA-3'	This study	N/A
v2 sgRNA scaffold F: 5'-GTTTTAGAGCGAGAAATCGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTGGATCCGCAACGGA-3'	This study	N/A
v2 sgRNA scaffold R: 5'-GAATCCGTTGCGGATCCAAAAAGCAC GACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTA ACTTGCATTCTCGCTCTA-3'	This study	N/A
SA157: 5'-GGATCCGCAACGGAATTC-3'	This study	N/A
SA158: 5'-GGTTGCGTCAGCAAACACAG-3'	This study	N/A
SA159: 5'-CAAGCAGAAGACGGCATAACGAGATGTGACTGGAGTT CAGACGTGTGCTCTCCGATCTGGTTCGTCAGCAAACACAG-3'	This study	N/A
Index primer SA160: 5'-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTCCGATCTTCCAGGCATGGATCCG CAACGGAATTC-3'	This study	N/A
Index primer SA161: 5'-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTCCGATCTATGCAGCTACGGATCC GCAACGGAATTC-3'	This study	N/A
Index primer SA162: 5'-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTCCGATCTGATCAATGTTCCGGATCC GCAACGGAATTC-3'	This study	N/A
Index primer SA163: 5'-AATGATACGGCGACCACCGAGATCTACA CTCTTTCCCTACACGACGCTCTCCGATCTCGATCTGGCGAAGG ATCCGCAACGGAATTC-3'	This study	N/A
Index primer SA164: 5'-AATGATACGGCGACCACCGAGATCTACA CTCTTTCCCTACACGACGCTCTCCGATCTTCTGTTCTGGGATCC GCAACGGAATTC-3'	This study	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Index primer SA165: 5'-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTTCCGATCTATCAGAACATGGATCC GCAACGGAATTC-3'	This study	N/A
Index primer SA166: 5'-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTTCCGATCTGATGCGATTGCGGATC CGCAACGGAATTC-3'	This study	N/A
Index primer SA167: 5'-AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTTCCGATCTCGATTTTGAGGTGGATC CGCAACGGAATTC-3'	This study	N/A
SA60: 5'-AATGATACGGCGACCACCGA-3'	This study	N/A
SA61: 5'-CAAGCAGAAGACGGCATACGA-3'	This study	N/A
<b>Recombinant DNA</b>		
pAWp28	Addgene	#73850
pAWp92	Addgene	#157975
pAWp100	Addgene	#157977
pAWp12	Addgene	#72732
pCMV-VSV-G	Addgene	#8454
pCMV-dR8.2 dvpr	Addgene	#8455
pAWp30	Addgene	#73857
pAWp63-clone32 (Opti-SpCas9)	Addgene	#131736
<b>Software and algorithms</b>		
MatLab R2018b	MathWorks	N/A
CombiPIPE	This study	<a href="https://github.com/AWHKU/CombiPIPE">https://github.com/AWHKU/CombiPIPE</a>
<b>Other</b>		
0.45 µm filter	Pall	Cat#4614
Amicon Ultra-15 centrifugal filter unit	Millipore	Cat#UFC910096
BD LSR Fortessa flow cytometer	BD Biosciences	N/A
BD Influx cell sorter	BD Biosciences	N/A
1.5 mL DNA LoBind tubes	Eppendorf	Cat#022431021
DynaMag- spin magnet	Invitrogen	Cat#12320D
Agilent 2100 bioanalyzer	Agilent	N/A

**STEP-BY-STEP METHOD DETAILS**

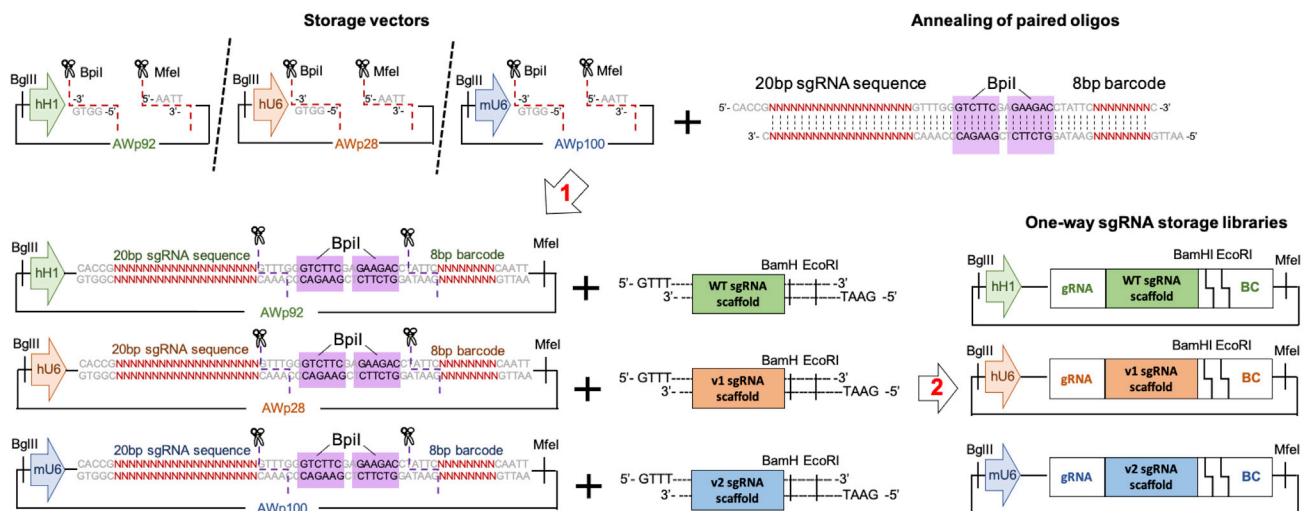
**Construction of barcoded single sgRNA storage vector libraries**

⌚ Timing: 1–2 weeks

This section describes the stepwise insertion of annealed barcoded sgRNA spacer-containing oligos and WT/v1/v2 sgRNA scaffold variants into storage vectors that contain engineered hH1/hU6/mU6 promoters to drive the expression of individual sgRNAs (Figure 1).

1. Anneal the oligo pairs.
  - a. Reconstitute each oligo to 25 µM with UltraPure H<sub>2</sub>O.
  - b. Anneal the forward and reverse oligo pairs individually by the following recipe:

Reagent	Volume (µL)
Forward oligo 1 (25 µM)	1
Reverse oligo 1 (25 µM)	1
10× Ligase buffer	1
T4 polynucleotide kinase (PNK)	0.5
UltraPure H <sub>2</sub> O	6.5
<b>Total</b>	<b>10 µL</b>



**Figure 1. Workflow for constructing barcoded single sgRNA storage vector libraries**

Incubate the mixture at 37°C for 30 min for T4 PNK to phosphorylate the 5'-hydroxyl terminus of each oligo, then anneal the oligo pair by denaturing at 98°C for 5 min and ramping down to 25°C at a rate of 1°C/s with a thermocycler.

- c. Pool 1  $\mu\text{L}$  of each annealed product, then dilute 5  $\mu\text{L}$  of the mixture by 50-fold with UltraPure H<sub>2</sub>O for the following ligation reaction in step 3. Store the rest of the annealed products at  $-20^{\circ}\text{C}$ .

**Note:** In (Zhou et al., 2020), 159 pairs of sgRNA spacer-containing oligos were annealed to give a pairwise sgRNA library which consists 25,281 combinations (159 X 159 sgRNAs). In building the three-way library, 32,768 combinations (32 X 32 X 32 sgRNAs) were built from 32 pairs of sgRNA spacer-containing oligos. The oligos can be synthesized and delivered in a 96- or 384-well microplate format, and multi-channel pipettes can be used to increase the handling throughput without robotic assistance. Alternatively, a customized size of pooled oligos can be synthesized on a chip for cloning into the storage vectors.

2. Digest storage vectors, pAWp28 (Addgene: #73850), pAWp92 (Addgene: #157975), and pAWp100 (Addgene: #157977) by Bpil and MfeI.
  - a. Prepare the following reaction for each storage vector:

Reagent	Volume ( $\mu\text{L}$ )
Storage vector (1 $\mu\text{g}$ )	4
10x FastDigest Green Buffer	2
Bpil	0.5
MfeI	0.5
UltraPure H <sub>2</sub> O	13
<b>Total</b>	<b>20 <math>\mu\text{L}</math></b>

- b. Incubate at 37°C for 15 min.
- c. Separate the digestion products by electrophoresis using 1% agarose gel stained with SYBR Safe DNA Gel Stain.

**Note:** The sizes of the digested fragments from pAWp28 are 4,231 bp, 113 bp, and 42 bp; from pAWp92 are 4,207 bp, 113 bp, and 36 bp; and from pAWp100 are 4,296 bp, 113 bp, and 26 bp.

- d. Extract the > 4 kb vector fragments by gel extraction kit, and elute the products in 20  $\mu$ L UltraPure H<sub>2</sub>O.
3. Ligate the annealed sgRNA spacer-containing oligos into the storage vectors.
    - a. Prepare the following reaction:

Reagent	Volume ( $\mu$ L)
Vector fragment (50 ng) (from step 2)	4
50-fold diluted annealed product mixture (from step 1)	1
10 $\times$ Ligase buffer	1
T4 DNA ligase	1
UltraPure H <sub>2</sub> O	3
<b>Total</b>	<b>10 <math>\mu</math>L</b>

- b. Incubate at 16°C for 16–18 h, then store the products at 4°C.

**Note:** The 3' end of the promoter sequences of hH1 and mU6 were modified to give the same compatible sticky ends such that the same annealed pool of gRNA oligos (from step 1) can be ligated to the three different vector fragments (from step 2) independently.

4. Transform the ligation products into Agilent X10-Gold Ultracompetent cells following manufacturer's instruction.

**Note:** After the recovery step of the transformation, we recommend taking small diluted portions of the total volume to spread on a LB-ampicillin agar plate, inoculate the rest into LB-ampicillin medium, and grow both the plate and liquid culture at 32°C for 16–20 h.

5. Quantify and sequence-verify the transformants.
  - a. Count the colony number on the plate, then the transformants in liquid culture could be estimated by multiplying the colony number by the proportion, e.g., number of colonies on agar plate  $\times$  dilution factor = number of transformants.
  - b. Extract plasmid DNA from the liquid culture by plasmid purification kit such as QIAGEN Plasmid Midi or Maxi kit.
  - c. Pick multiple colonies on the plate to be verified by Sanger sequencing.

**△ CRITICAL:** We recommend generating transformants with at least 80-fold representation of the library size. Repeat ligation until sufficient fold representation is reached. Combine the plasmids from different batches by the molar ratio of 1:1 to achieve higher library coverage without introducing bias.

6. Prepare vector fragments of single sgRNA spacer-containing storage vector libraries.
  - a. Digest each of the single sgRNA spacer-containing storage vector libraries generated in step 5 by Bpil:

Reagent	Volume ( $\mu$ L)
Single sgRNA storage vector (1 $\mu$ g)	2
10 $\times$ FastDigest Green Buffer	2
Bpil	0.5
UltraPure H <sub>2</sub> O	15.5
<b>Total</b>	<b>20 <math>\mu</math>L</b>

- b. Incubate at 37°C for 15 min.

- c. Separate the digested products by electrophoresis using 1% agarose gel with SYBR Safe DNA Gel Stain.
  - d. Extract the > 4 kb vector fragments from gel slice by gel extraction kit, and elute in 20  $\mu$ L UltraPure H<sub>2</sub>O.
7. Ligate the linearized single sgRNA spacer-containing storage vector libraries with annealed oligos containing the wild-type/v1/v2 sgRNA scaffold sequences.
- a. Anneal oligos WT sgRNA scaffold F + WT sgRNA scaffold R, v1 sgRNA scaffold F + v1 sgRNA scaffold R, and v2 sgRNA scaffold F + v2 sgRNA scaffold R to prepare the wild-type, v1, and v2 sgRNA scaffold sequence inserts, respectively. The reaction is performed as described in step 1.
  - b. Prepare the following ligation reaction:
  - c. Incubate at 16°C for 16–18 h, then store the products at 4°C.

Reagent	Volume ( $\mu$ L)
Vector fragment (50 ng) (from step 6)	4
50-fold diluted annealed products (from (a))	1
10 $\times$ Ligase buffer	1
T4 DNA ligase	1
UltraPure H <sub>2</sub> O	3
<b>Total</b>	<b>10 <math>\mu</math>L</b>

**Note:** In the current design of CombiGEM-CRISPR v2.0, the annealed products harboring wild-type, v1, and v2 sgRNA scaffold sequence inserts are ligated to the digested vectors containing hH1, hU6, and mU6, respectively.

8. Transform the products into Agilent X10-Gold Ultracompetent cells as in step 4.
9. Quantify and sequence-verify the transformants as in step 5. Similarly, we recommend the total number of transformants collected to be at least 80-fold representation of the library size.

### Assembly of barcoded multiplexed sgRNA libraries

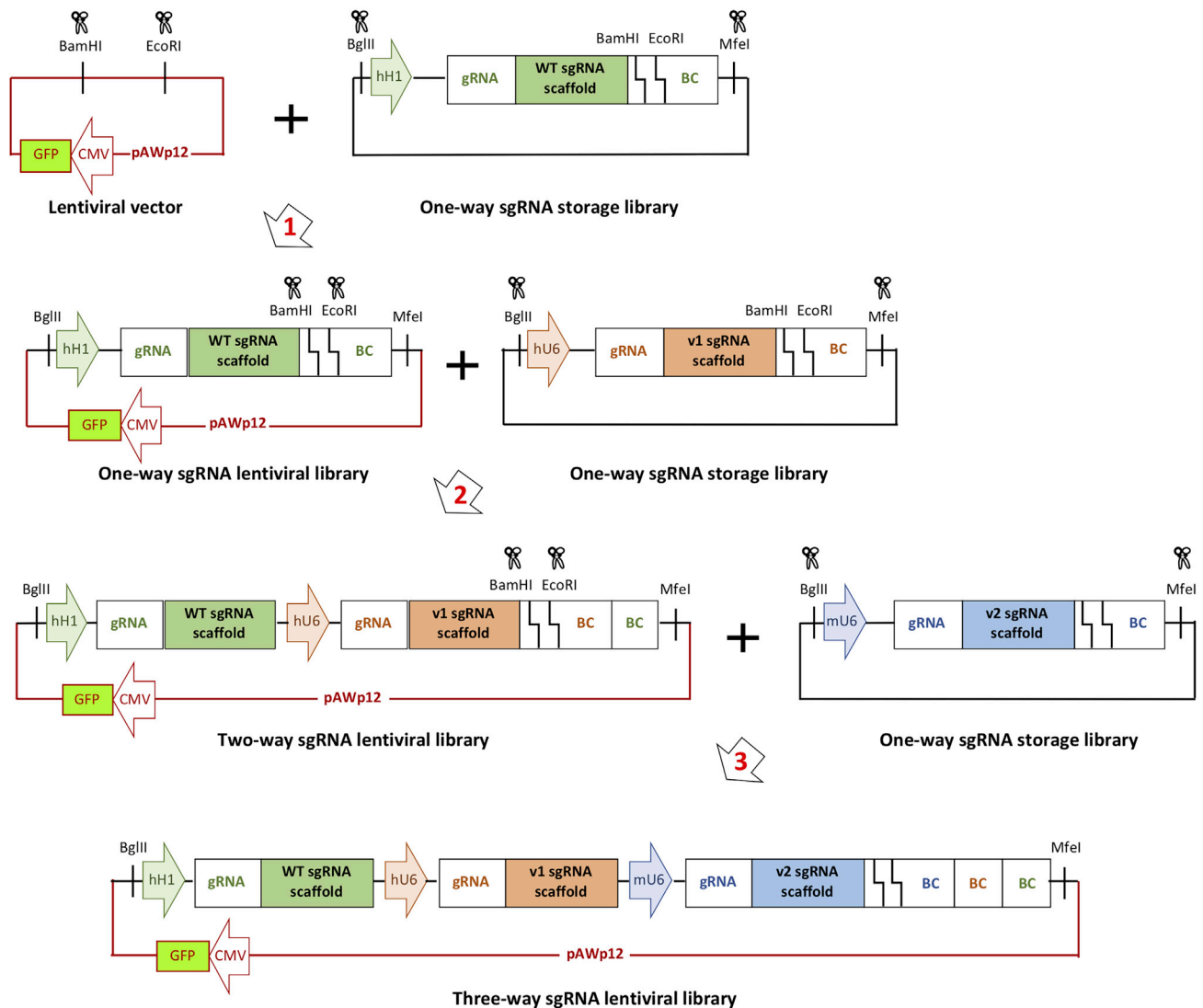
⌚ **Timing: 2–3 weeks**

This section provides detailed procedures to build the two- and three-way sgRNA libraries. Steps 10–14 describes the generation of the one-way lentiviral vector-based sgRNA-expressing library; step 15 describes the assembly of the two-way sgRNA library; and step 16 describes the assembly of the three-way sgRNA library (Figure 2). The multiplexed sgRNA expression cassettes are built with combinations of sgRNA scaffold variants and promoters, including human H1, human U6, and mouse U6, that have been reported to efficiently express sgRNAs in mammalian cells (Adamson et al., 2016). In (Zhou et al., 2020), these promoter and scaffold variants were tested to give similar gene knockout efficiency, suggesting that there was no promoter usage bias and that similar amounts of sgRNAs were generated to evenly knock out individual genes.

10. Prepare the insert fragments from the single sgRNA storage vector libraries.
  - a. Digest each of the single sgRNA storage vector libraries (from step 9) by BglIII and MfeI.

Reagent	Volume ( $\mu$ L)
sgRNA storage vector library (1 $\mu$ g)	2
10 $\times$ FastDigest Green Buffer	2
BglIII	0.5
MfeI	0.5
UltraPure H <sub>2</sub> O	15
<b>Total</b>	<b>20 <math>\mu</math>L</b>





**Figure 2. Workflow for assembling barcoded multiplexed sgRNA libraries**

- b. Incubate at 37°C for 15 min.
- c. Separate the fragments by electrophoresis using 1% agarose gel with SYBR Safe DNA Gel Stain.
- d. Extract the ~400 bp fragments (see the Note below) from gel slice by gel extraction kit, then elute in 20  $\mu$ L UltraPure H<sub>2</sub>O and measure DNA concentration by NanoDrop spectrometer.

**Note:** The insert fragment harboring H1 promoter and wildtype sgRNA scaffold together with the sgRNA library (stored in pAWp92) has a size of 365 bp; the insert fragment harboring U6 promoter and v1 sgRNA scaffold together with the sgRNA library (stored in pAWp28) has a size of 389 bp; and the insert fragment harboring mU6 promoter and v2 sgRNA scaffold together with the sgRNA library (stored in pAWp100) has a size of 454 bp.

11. Prepare the lentiviral vector.
  - a. Digest the destination lentiviral backbone, pAWp12 (Addgene: #72732), by BamHI and EcoRI.

Reagent	Volume (μL)
pAWp12 (1 μg)	2
10× FastDigest Green Buffer	2
BamHI	0.5
EcoRI	0.5
UltraPure H <sub>2</sub> O	15
<b>Total</b>	<b>20 μL</b>

- b. Incubate at 37°C for 15 min.
- c. Separate the fragments electrophoresis using 1% agarose gel with SYBR Safe DNA Gel Stain.
- d. Extract the 9.3 kb vector fragments from gel slice by gel extraction kit, then elute in 20 μL UltraPure H<sub>2</sub>O and measure DNA concentration by NanoDrop spectrometer.

12. Ligate the single sgRNA expression cassette into lentiviral vector to assemble the one-way lentiviral vector-based sgRNA-expressing library
  - a. Prepare the following ligation reaction:

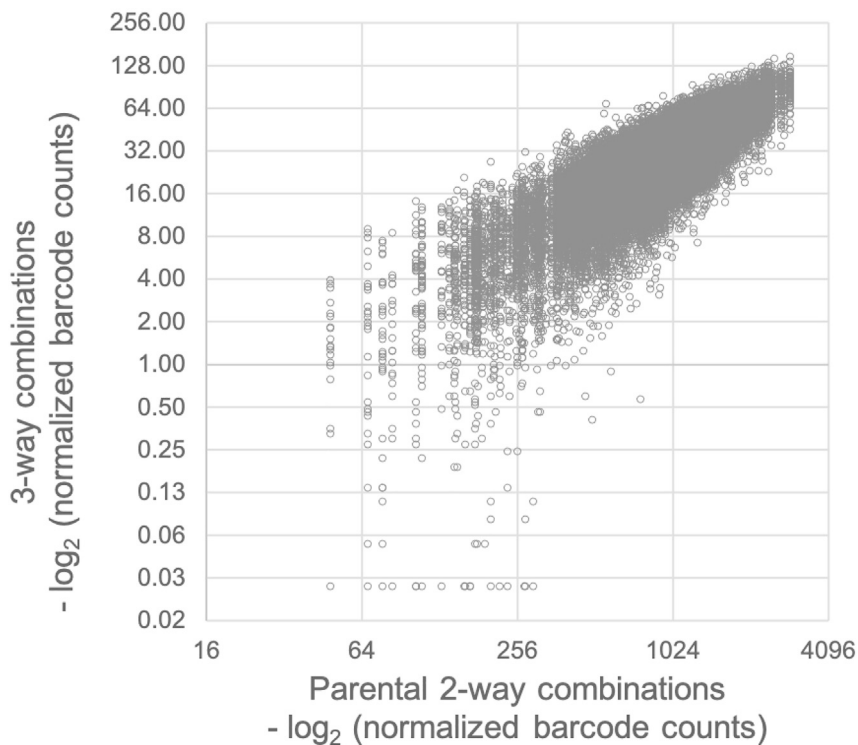
Reagent	Volume (μL)
Digested pAWp12 (50 ng) (from step 11)	4
Insert fragment stored in pAWp92 (5 ng) (from step 10)	1
10× Ligase buffer	1
T4 DNA ligase	1
UltraPure H <sub>2</sub> O	3
<b>Total</b>	<b>10 μL</b>

- b. Incubate at 16°C for 16–18 h, then store the products at 4°C.

13. Transform the products into Agilent XL10-Gold Ultracompetent cells as in step 4.
14. Collect sufficient number of transformants and sequence-verify them as in step 5.
15. To build the two-way sgRNA library, repeat steps 11–14 with two adjustments: (i) using the insert fragment stored in pAWp28 instead of pAWp92 in step 10; (ii) using the one-way lentiviral vector-based sgRNA-expressing library (from step 14) instead of pAWp12 in step 11.
16. To build the three-way sgRNA library, repeat steps 11–14 with two adjustments: (i) using the insert fragment stored in pAWp100 instead of pAWp92 in step 10; (ii) using the two-way lentiviral vector-based sgRNA-expressing library (from step 15) instead of pAWp12 in step 11.

**△ CRITICAL:** Again, we recommend generating transformants with at least 80-fold representation of the library size for each of the above libraries. Repeat ligation until sufficient fold representation is reached.

**Note:** In addition to sequence-verify the transformants, restriction digestion reactions could also be used to check for expected band sizes excised from the correctly assembled one-, two-, and three- way plasmid library pools. The expected band sizes after digestion by BglII and EcoRI are ~9.3 kb plus ~400 bp (for the one-way library), ~1 kb (for the two-way library), or ~1.5 kb (for the three-way library). With the every additional step in the cloning process, it is anticipated that some of the library members may be depleted. While the library coverage of the three-way sgRNA pool was 99.8% as revealed by sequencing, comparison of sgRNA composition between the three-way sgRNA plasmid library products (harboring 32,768 total sgRNA combinations described in (Zhou et al., 2020)) and their corresponding parental two-way sgRNA plasmid library products revealed that three-way combinations of which their parental two-way constructs had lower abundance in the pool could be more prone to depletion (Figure 3). Overall, only a small number of library members are depleted or lost in the process.



**Figure 3. Comparison of sgRNA composition between the pools of three-way sgRNA plasmid library products and their parent two-way sgRNA library products**

**Note:** Other orderings of promoters and sgRNA scaffold variants are possible while those configurations should be tested for efficiency before the assembly of combinatorial sgRNA libraries.

### Delivery of multiplexed sgRNA library for screening in mammalian cells

⌚ Timing: 2–4 weeks (varies among the screening phenotypes)

This section outlines the steps for preparing lentiviral particles to transduce the mammalian cells of interest with the sgRNA library.

17. Produce lentiviral particles for delivering the sgRNA library into mammalian cells that stably express *Streptococcus pyogenes* Cas9 (SpCas9).

**Note:** For packaging the multiplexed sgRNA library assembled in pAWp12, 2<sup>nd</sup> generation packaging plasmids (i.e., pCMV-dR8.2 dvpr [Addgene: #8455] and pCMV-VSV-G [Addgene: #8454]) can be used. 3<sup>rd</sup> generation packaging plasmids can also be used. A detailed protocol for lentivirus production was described in (Brown et al., 2020). The lentiviral-based plasmids that express wild-type SpCas9 (Addgene: #73857) and high-fidelity SpCas9 such as Opti-SpCas9 (Addgene: #131736) could be used for generating the stable SpCas9-expressing cell line.

18. Estimate the titer of the lentiviral particles.

- a. Seed mammalian cells onto a 24-well cell culture plate at a confluence of 30% in 0.34 mL complete culture medium.
- b. Add a serial volume of viral medium (e.g., 10, 40, 80, and 160  $\mu$ L) into each well, and top up the total volume of each well to 0.5 mL by complete culture medium. Leave one well without virus as negative control.
- c. Replace medium of each transduced well after 24 h.
- d. Incubate the cells for another 4–5 days.
- e. Determine the GFP-positive cell population under each condition by flow cytometry.

**Note:** The amount of viral medium needed for efficiently infecting the cells of interest varies among different cell types, and it has to be empirically determined if prior information is not available. For some difficult-to-infect cell types, the viral medium can be concentrated using a filtration unit such as the Amicon Centrifugal filter. The green fluorescent protein (GFP) expressed from the lentiviral vector pAWp12 serves as a transduction marker. The virus amount for approximately 30% transduction, which infers a multiplicity of infection (MOI) of 0.3–0.5, could be estimated from the data plot.

19. Transduce mammalian cells with the lentiviral particles.
  - a. Seed the required number of cells at a confluence of 30% in complete culture medium.
  - b. Add viruses to transduce cells at a MOI of 0.3–0.5. The virus volume needed can be determined from step 18 and scaled up proportionally based on the number of cells needed for the screening.
  - c. Change medium of transduced cells after 24 h.
  - d. Maintain the cells for another 4–5 days with confluence lower than 80%.

**△ CRITICAL:** The number of cells being transduced should reach at least 300-fold representation of the library size. A higher fold representation (e.g., 3,000-fold) is recommended if cells are not limited.

20. Perform cell sorting by flow cytometry to enrich infected cells with GFP signal.

**Note:** The actual transduction rate can be determined at this point.

**△ CRITICAL:** The number of cells being sorted into the GFP-positive gate should reach at least 100-fold representation of the library size. Again, a higher fold representation (e.g., 1,000-fold) is recommended if cells are not limited. It is recommended to take at least 100-fold representation of the sorted cells for genomic DNA extraction as the initial sample of cell library. We recommend collecting the initial cell library sample as early as possible, to minimize any gene knockout effect that could be introduced to the library distribution.

21. Screen the cell library based on the phenotypes of interest. For example, to identify the combinatorial effects of double- or triple-gene knockout on cell proliferation, the cell library can be propagated for a certain period. Collect genomic DNA from the cells at different time points. Compare the change in the barcode abundance between the later time points and the initial time point to reveal combinations with enrichment or depletion.

**△ CRITICAL:** During the propagation of the cell pool, cell confluency should not exceed 80%. If cell passaging is needed during the course of an assay, cell number of no less than 100-fold representation of the library size should be passaged to the new culture dish. Again, a higher fold representation (e.g., 1,000-fold) is recommended. If multiple culture dishes are used, cells from all dishes after trypsinization should be combined and mixed before taking sufficient cells to new dishes.

### Sample preparation for barcode sequencing

⌚ Timing: 1 day

This section describes the steps to extract barcodes from the plasmid and infected cell libraries for high-throughput sequencing.

22. Extract genomic DNA from cells using the QIAGEN DNasey Blood and Tissue Kit, following the manufacturer's instructions. Measure the DNA concentration by NanoDrop spectrometer.
23. PCR amplification of the barcode region from genomic DNA and plasmid DNA library samples.
  - a. Prepare the following PCR reaction for genomic DNA template.

Reagent per one PCR reaction tube	Volume (μL)
Genomic DNA (800 ng)	4
SA157 (20 μM)	0.75
SA158 (20 μM)	0.75
KAPA HiFi RM (2x)	25
UltraPure H <sub>2</sub> O	19.5
<b>Total</b>	<b>50 μL</b>

- b. Prepare the following PCR reaction for plasmid DNA template.

Reagent	Volume (μL)
Plasmid library (0.5 ng)	1
SA157 (20 μM)	0.2
SA158 (20 μM)	0.2
KAPA HiFi RM (2x)	6.25
UltraPure H <sub>2</sub> O	4.85
<b>Total</b>	<b>12.5 μL</b>

- c. Run the PCR according to the following cycling conditions.

PCR Cycling Conditions			
Steps	Temperature	Time	Cycles
Initial Denaturation	95°C	5 min	1
Denaturation	98°C	20 s	18 cycles (for genomic DNA)
Annealing	60°C	15 s	14 cycles (for plasmid)
Extension	72°C	20 s	
Final extension	72°C	20 s	1
Hold	16°C		

The expected band sizes should be 298 bp for two-way library and 312 bp for three-way library.

⚠ **CRITICAL:** The optimal PCR cycle number was determined by qPCR to ensure the amplification lies in the exponential phase (Wong et al., 2015; Wong et al., 2016).

⚠ **CRITICAL:** We do not recommend exceeding 800 ng of genomic DNA per 50 μL reaction; therefore, to cover the library by at least 100-fold representation, one needs to calculate how many PCR reactions are required. For example, one copy of the normal human genome is estimated to weigh  $6.5 \times 10^{-12}$  g. Given a library size of 30,000 combinations,

the genomic DNA input for 100-fold representation would be approximately  $1.95 \times 10^4$  ng, which requires 25 reactions. 0.5 ng of the plasmid DNA library is estimated to contain approximately  $3 \times 10^7$  copies of library elements; therefore, one reaction is enough.

24. For genomic DNA samples, select and purify the PCR products of expected size using AMPure XP beads (0.95–0.5×) following the manufacturer’s instruction.
  - a. Combine all PCR products for each sample and mix by low-speed vortexing.
  - b. Take 200  $\mu$ L of the PCR products to mix with 100  $\mu$ L (0.5×) of resuspended AMPure XP beads in a 1.5-mL tube.
  - c. Incubate the mixture at 20°C–25°C for 5 min, then put the tube onto a magnetic stand, such as DynaMag- Spin Magnet (Invitrogen).
  - d. Transfer the 300  $\mu$ L of supernatant into a new 1.5-mL LoBind tube and mix with 90  $\mu$ L (0.95×) of resuspended AMPure XP beads.
  - e. Incubate the mixture at 20°C–25°C for 5 min, then put the tube onto a magnetic stand.
  - f. Remove the supernatant carefully by pipetting and wash beads with 500  $\mu$ L of freshly prepared 85% ethanol once.
  - g. Aspirate the supernatant and allow the beads to air-dry for 1–2 min.
  - h. Elute the products in 41  $\mu$ L UltraPure H<sub>2</sub>O.
  - i. Transfer the supernatant to a clean PCR tube.

**△ CRITICAL:** Sufficient amount of PCR product should be applied for purification. We use 200  $\mu$ L of PCR product for each purification and size selection in one reaction tube.

25. To generate next-generation sequencing-compatible and indexed libraries for sequencing, incorporate adapters and indexes to amplicons during library amplification.
  - a. Prepare the following PCR reaction for genomic DNA amplicons (from step 24).

Reagent per reaction	Volume ( $\mu$ L)
Purified genomic DNA amplicons	10
SA159 (20 $\mu$ M)	0.75
Index primer (20 $\mu$ M)	0.75
KAPA HiFi RM (2×)	25
UltraPure H <sub>2</sub> O	13.5
<b>Total</b>	<b>50 <math>\mu</math>L</b>

- b. Prepare the following PCR reaction for plasmid DNA amplicons (from step 23).

Reagent	Volume ( $\mu$ L)
Unpurified plasmid amplicons	5
SA159 (20 $\mu$ M)	0.75
Index primer (20 $\mu$ M)	0.75
KAPA HiFi RM (2×)	25
UltraPure H <sub>2</sub> O	18.5
<b>Total</b>	<b>50 <math>\mu</math>L</b>

- c. Run the PCR according to the following cycling conditions.

PCR Cycling Conditions			
Steps	Temperature	Time	Cycles
Initial Denaturation	95°C	5 min	1
Denaturation	98°C	20 s	8 – 10 cycles
Annealing	60°C	15 s	
Extension	72°C	20 s	
Final extension	72°C	20 s	1
Hold	16°C		

The expected band sizes should be 423 bp for two-way library and 437 bp for three-way library (1–4 bp variation in length of the products according to the index primer used).

**△ CRITICAL:** To minimize any bias introduced by the PCR, the cycle number should be maintained as low as possible.

26. Purify the PCR products from plasmid and cell libraries (from step 25) by AMPure XP beads (0.95–0.5x) as described in step 25 and with reaction volumes scaled down accordingly. Elute the products in 11  $\mu\text{L}$  of UltraPure  $\text{H}_2\text{O}$ .
27. Quantification of the resultant barcode-containing libraries (from step 26) by qPCR (NEB kit and SYBR Green PCR Mix).
  - a. Prepare serial dilutions using 1  $\mu\text{L}$  of the libraries (e.g., from 1,000- to 20,000-fold) using UltraPure  $\text{H}_2\text{O}$ . One of the serial dilutions should fall within the dynamic range of the DNA standards.
  - b. Prepare the following PCR reaction for library samples.

Reagent	Volume ( $\mu\text{L}$ )
Diluted library	1
SA60 (20 $\mu\text{M}$ )	0.2
SA61 (20 $\mu\text{M}$ )	0.2
2x SYBR premix	5
UltraPure $\text{H}_2\text{O}$	3.6
<b>Total</b>	<b>10 <math>\mu\text{L}</math></b>

- c. Prepare the following PCR reaction for standards.

Reagent	Volume ( $\mu\text{L}$ )
NEB standard	0.5
SA60 (20 $\mu\text{M}$ )	0.2
SA61 (20 $\mu\text{M}$ )	0.2
2x SYBR premix	5
UltraPure $\text{H}_2\text{O}$	4.1
<b>Total</b>	<b>10 <math>\mu\text{L}</math></b>

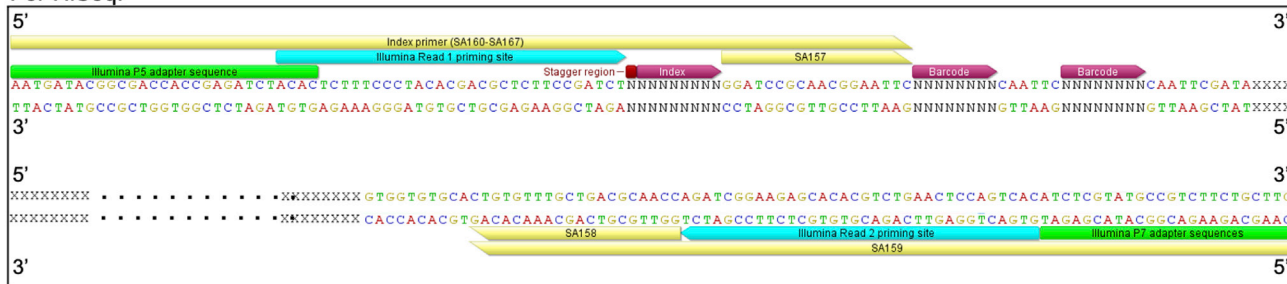
- d. Run the reactions on qPCR machine to obtain  $C_T$  values, and calculate the molar concentration accordingly.
28. Pool the indexed libraries at equal molar ratio to a final concentration of 20 nM in approximately 20  $\mu\text{L}$ , and run 1  $\mu\text{L}$  of purified products on the Bioanalyzer. Check peak distribution.
29. Run the sample using Illumina HiSeq.

**Note:** To help with sequencing low-complexity libraries, 10%–30% Phi-X can be spiked in. Alternatively, staggered primer design (i.e., with stagger regions of different lengths in the index primers) can be introduced to maintain sequence diversity. Other high-throughput sequencing platforms such as NovaSeq could be used and the adaptor sequences may need to be modified accordingly. The architecture of the sequencing the barcodes for a pairwise sgRNA library and reads using HiSeq and NovaSeq is shown in [Figure 4](#) as an example.

### Sequencing data analysis using CombiPIPE

⌚ **Timing:** variable (typically up to 40 min for an 8 Gb fastq file, depending on the computer used for the analysis, size of the library, and sequencing depth; ~1 min for the test dataset provided on GitHub)

For HiSeq:



For NovaSeq:

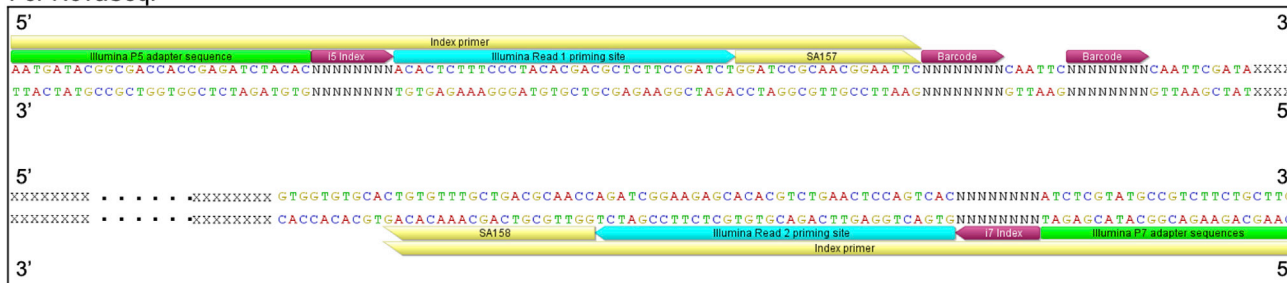


Figure 4. Architecture of sequencing libraries

This section describes a step-by-step walkthrough of CombiPIPE, an analysis pipeline containing the steps for processing raw sequencing data generated from the CombiGEM-CRISPR v2.0 platform and outputting files with the following information for each combination calculated: counts per million reads (CPM),  $\log_2$  fold change ( $\log_2FC$ ),  $-\log$  p-value ( $-\log_{10}PV$ ), genetic interaction (GI), Dunnett test adjusted p-value (Figure 5). The CombiPIPE pipeline can be accessed at <https://github.com/awhku/combiPIPE>. The pipeline is written in nextflow, and running the pipeline requires nextflow installation (<https://nf-co.re/usage/installation>). In addition, CombiPIPE depends on python3 and requires the python packages regex (<https://pypi.org/project/regex/>), scipy (<https://pypi.org/project/scipy/>), and numpy (<https://pypi.org/project/numpy/>), and the R package DescTools (<https://www.rdocumentation.org/packages/DescTools/versions/0.99.37>) to be installed to run CombiPIPE successfully:

- The bin directory contains the scripts in python and R for each process in CombiPIPE:
  - process1\_sortSamples.py sorts the FASTQ entries to their according samples.
  - process2\_newBCanalyzer.py maps the FASTQ entries in each sample to the barcodes representing the gene targets.
  - process3\_calcFCPV.py calculates  $\log_2FC$  and  $-\log_{10}PV$  for each combination and filters out combinations with repeated gRNAs.
  - process4\_calcGI.py calculates the GI value by comparing the expected  $\log_2FC$ , the sum of the corresponding single gene knockouts, and the observed  $\log_2FC$ .
  - process5\_DunnettTest.R takes in the Dunnett test inputs and performs Dunnett tests on combinations included in FCPV.py.
  - process6\_mergeDunnetts.py merges the Dunnett test adjusted p-values from process 5.
- The nextflow.config file contains all arguments used when running CombiPIPE.
  - `-fastq`: path to the FASTQ raw sequence file generated using the CombiGEM-CRISPR v2.0 platform.
  - `-sampleinfo`: a CSV file with three columns without headers:

Sample Sequence Tag	Sample Name	Sample Condition
---------------------	-------------	------------------



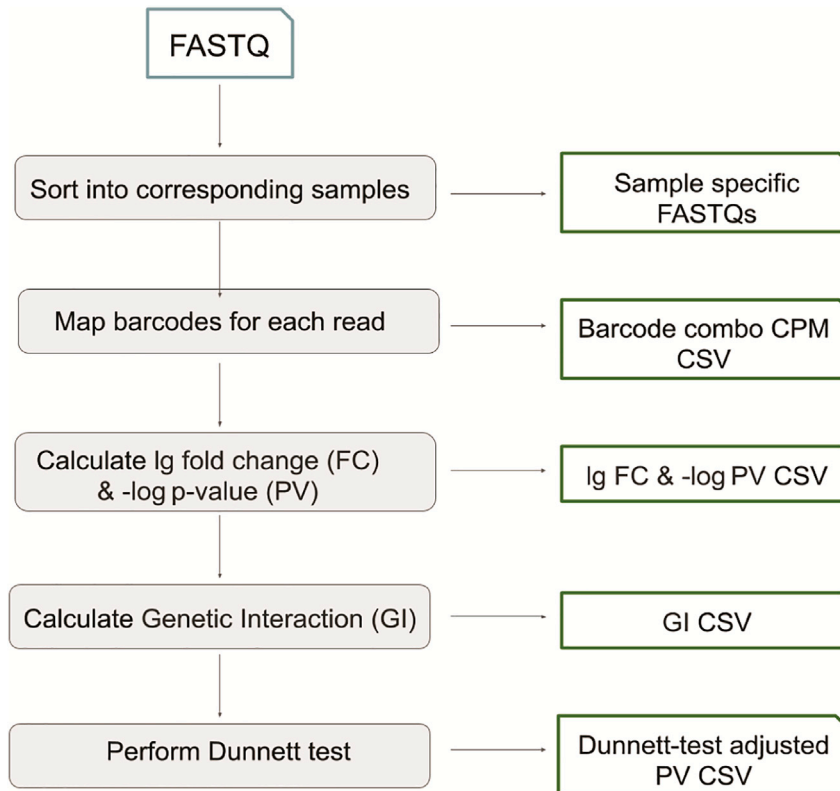


Figure 5. Workflow of CombiPIPE

- `-barcodes`: a CSV file with three columns without headers:

Barcode	Number Tag	gRNA target
---------	------------	-------------

- `-pattern`: the first 7 letters in the first line of each FASTQ entry.
- `-dimension`: the combination dimension.
- `-linker`: the linker sequence connecting the barcodes.
- `-dummysgs`: the number tags for the dummy sgRNA controls (default: 1,2).
- The `CombiGEM.nf` file is the nextflow pipeline that connects the processes in CombiPIPE.

**Note:** For demonstration of the pipeline and testing purposes, a minimal test dataset was generated from the three-way gene knockout screen described in (Zhou et al., 2020) and can be found under the `test-dataset` directory. This directory contains a compressed fastq file and the corresponding sample information file (`sampleInfo.csv`) and barcode information file (`barcode_list.csv`) required for sorting samples in the library and mapping the raw sequence data to the barcodes representing the gene targets.

- Clone or download the pipeline in your project directory using the link reported on GitHub:  
\$ git clone <https://github.com/AWHKU/CombiPIPE.git>
- Get into the CombiPIPE directory:  
\$ cd CombiPIPE

```
(base) aws-iMac:CombiPIPE-master awlab$ nextflow CombiGEM.nf --fastq test-dataset/OVARIAN_testdataset.fastq --sampinfo test-dataset/sampleInfo.csv --barcodes test-dataset/barcode_list.csv --pattern @D00691 --dimensions 3 --linker CAATTC --dummysgs 1,2
N E X T F L O W ~ version 20.01.0
Launching `CombiGEM.nf` [clever_brahmagupta] - revision: d62840b891
executor > local (6)
[d6/9df9fb] process > sortSamples [100%] 1 of 1 ✓
[27/a42228] process > newBCanalyzer [100%] 1 of 1 ✓
[22/ecf5bb] process > log2FC_neglogPV [100%] 1 of 1 ✓
[f3/70a600] process > PairwiseGI_genDunnettInputs [100%] 1 of 1 ✓
[90/fe8836] process > batchDunnettTest [100%] 1 of 1 ✓
[19/8cde22] process > mergeDunnetts [100%] 1 of 1 ✓
Completed at: 18-Aug-2020 16:10:23
Duration : 1m
CPU hours : (a few seconds)
Succeeded : 6
```

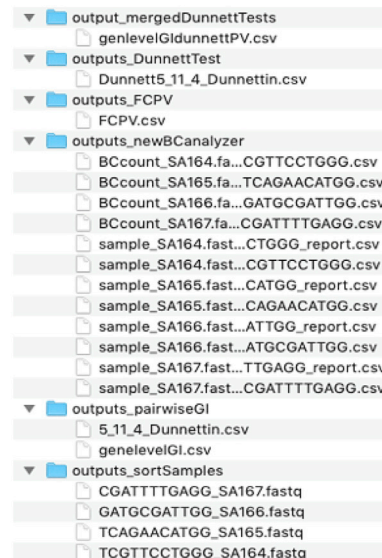


Figure 6. Sample output after CombiPIPE analysis

32. Uncompress the fastq file in the test-dataset directory:  
\$ gunzip test-dataset/OVARIAN\_testdataset.fastq.gz
33. Run the pipeline with nextflow:  
\$ nextflow CombiGEM.nf --fastq test-dataset/OVARIAN\_testdataset.fastq --sampinfo test-dataset/sampleInfo.csv --barcodes test-dataset/barcode\_list.csv --pattern @D00691 --dimensions 3 --linker CAATTC --dummysgs 1,2

### EXPECTED OUTCOMES

When CombiPIPE finishes running without errors, the following files should be generated (Figure 6):

```
├ outputs_sortSamples
│   └ <sample_sequence_tag>_<sample_name>.fastq
├ outputs_newBCanalyzer
│   ├── sample_<sample_name>.fastq_<sample_sequence_tag>.csv
│   ├── sample_<sample_name>.fastq_<sample_sequence_tag>_report.csv
│   └ BCcount_<sample_name>.fastq_<sample_sequence_tag>.csv
├ outputs_FCPV
│   └ FCPV.csv
├ outputs_pairwiseGI
│   ├── genelevelGI.csv
│   └ <combination_name>_Dunnettin.csv
├ outputs_DunnettTest
│   └ Dunnett<combination_name>_Dunnettin.csv
├ outputs_mergedDunnettTests
│   └ genlevelGIdunnettPV.csv
```

- outputs\_sortSamples: this directory contains the FASTQ files from samples indicated in the sample information CSV file.
- outputs\_newBCanalyzer: this directory contains the following files:
  - The sample\_<sample\_name>.fastq\_<sample\_sequence\_tag>.csv files contain the barcodes of the successfully mapped entries.
  - The sample\_<sample\_name>.fastq\_<sample\_sequence\_tag>\_report.csv files contain the counts for the total, successfully mapped, discarded reads

- The BCcount\_<sample\_name>.fastq\_<sample\_sequence\_tag>.csv files contain the counts and CPM for each of the barcode combinations.
- outputs\_FCPV: this directory contains the FCPV.csv file, which contains the observed  $\log_2FC$ s and  $-\log_{10}PV$ s.
- outputs\_pairwiseGI: this directory contains the following files:
  - The genelevelGI.csv file contains the expected  $\log_2FC$ s, observed  $\log_2FC$ s, GI values for all combinations in FCPV.csv.
  - The <combination\_name>\_Dunnett.csv file(s) as Dunnett test inputs.
- outputs\_DunnettTest: this directory contains all output files generated from running Dunnett test on each combination.
- outputs\_mergedDunnettTests: this directory contains the genelevelGIdunnettPV.csv file, which contains expected  $\log_2FC$ s, observed  $\log_2FC$ s, GI values, and Dunnett test adjusted p-values for all combinations in FCPV.csv.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Each read produced from the CombiGEM-CRISPR v2.0 platform is ordered from barcode<sub>n</sub> to barcode<sub>1</sub> due to how the combinatorial sgRNA library is assembled. Barcodes are mapped to each FASTQ entry with a maximum hamming distance of 1. Based on our experience, we observed an average of about 95.8% (ranged from 84.9% to 99.6%) of mappable barcodes as assessed by NGS. Counts of each barcode combination are normalized to CPM (counts per million reads).  $\log_2FC$  (fold change) and  $-\log_{10}PV$  (p-value) are calculated based on pre- and post-treatment sample replicates. GI (genetic interaction) value is calculated by *observed  $\log_2FC$  - expected  $\log_2FC$* , which is the sum of the corresponding single gene knockouts. The Dunnett test is performed on combinations without repeated sgRNAs by comparing the highest order of combinatorial knockout population to its respective lower order of combinatorial and/or single knockout populations.

## LIMITATIONS

This protocol is suitable for two- and three-way combinatorial sgRNA libraries. We do not recommend one to generate even higher-order combinations based on the reagents provided due to the possibility of homologues recombination of the vectors. However, if more sequence-diversified Pol III promoters and sgRNA scaffold variants could be employed, higher-order combinations are achievable according to this protocol. The packaging capacity of the lentiviral vector is also a limitation for one to consider when extending the combinatorial libraries with long DNA elements.

## TROUBLESHOOTING

### Problem 1

Transformation has efficiency lower than expected.

### Potential solution

Transformation efficiency could vary among batches of competent cells. Avoid freeze-and-thaw cycles of competent cells to maintain their competence. Other chemically competent or electrocompetent cells can also be used for transformation.

### Problem 2

Sanger sequencing of colonies from the pool of transformants identifies original vector plasmid sequence with no insert.

### Potential solution

The restriction digestion reactions have to be performed at their optimal buffered conditions to ensure complete digestion of the destination vectors, since any residual undigested vectors could result in its amplification after transformation. Our experience with the above listed restriction enzymes can result in very efficient digestions.

### Problem 3

Insufficient PCR products after library sample preparation for barcode sequencing.

### Potential solution

Add minimal number of PCR cycles in step 25.

### Problem 4

The restriction enzymes used in the library construction by the CombiGEM-CRISPR v2.0 approach are not compatible with other destination vectors required for your specific cell systems because they carry those restriction sites.

### Potential solution

Other enzyme pairs that give compatible ends for ligation can be used for CombiGEM-CRISPR v2.0. It is best to choose enzymes that can cut efficiently in a single buffer for a double digestion. We do not recommend doing sequential digestion that has a lower recovery rate of digested vectors/inserts, since it is harder to maintain the representation of library members during the cloning process.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Alan S.L. Wong ([aslw@hku.hk](mailto:aslw@hku.hk)).

### Materials availability

Plasmids described have been deposited to Addgene.

### Data and code availability

The pipeline and codes for analyzing CombiGEM-CRISPR v2.0 data are available on <https://github.com/AWHKU/CombiPIPE>.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xpro.2020.100255>.

## ACKNOWLEDGMENTS

We thank members of the Wong lab for helpful discussions. This work was supported by the Hong Kong Research Grants Council (ECS-27105716, GRF-17102218, and TRS-T12-710/16-R), Innovation and Technology Commission (ITS/133/18), and Food and Health Bureau (HMRP-04151416) (to A.S.L.W.).

## AUTHOR CONTRIBUTIONS

P.Z., Y.K.W., B.K.C.C., G.C.G.C., and A.S.L.W. wrote the manuscript.

## DECLARATION OF INTERESTS

The authors have filed patent applications based on this system.

## REFERENCES

- Adamson, B., Norman, T.M., Jost, M., Cho, M.Y., Nunez, J.K., Chen, Y., Villalta, J.E., Gilbert, L.A., Horlbeck, M.A., Hein, M.Y., et al. (2016). A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* *167*, 1867–1882.e21.
- Allen, F., Crepaldi, L., Alsinet, C., Strong, A.J., Kleshchevnikov, V., De Angeli, P., Palenikova, P., Khodak, A., Kiselev, V., Kosicki, M., et al. (2018). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol.* <https://doi.org/10.1038/nbt.4317>.
- Brown, L.Y., Dong, W., and Kantor, B. (2020). An improved protocol for the production of lentiviral vectors. *STAR Protoc.* *1*, 100152.
- Choi, G.C.G., Zhou, P., Yuen, C.T.L., Chan, B.K.C., Xu, F., Bao, S., Chu, H.Y., Thean, D., Tan, K., Wong, K.H., et al. (2019). Combinatorial mutagenesis en masse optimizes the genome editing activities of SpCas9. *Nat. Methods* *16*, 722–730.
- Chuai, G.H., Wang, Q.L., and Liu, Q. (2017). In silico meets in vivo: towards computational CRISPR-based sgRNA design. *Trends Biotechnol.* *35*, 12–21.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191.

Perez, A.R., Pritykin, Y., Vidigal, J.A., Chhangawala, S., Zamparo, L., Leslie, C.S., and Ventura, A. (2017). GuideScan software for improved single and paired CRISPR guide RNA design. *Nat. Biotechnol.* 35, 347–349.

Perez, A.R., Sala, L., Perez, R.K., and Vidigal, J.A. (2019). Correction of off-targeting in CRISPR screens uncovers genetic dependencies in melanoma cells. *bioRxiv*, 809970.

Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E., Vaimberg, E.W., Goodale, A., Root, D.E., Piccioni, F., et al. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* 9, 5416.

Shen, M.W., Arbab, M., Hsu, J.Y., Worstell, D., Culbertson, S.J., Krabbe, O., Cassa, C.A., Liu, D.R., Gifford, D.K., and Sherwood, R.I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* 563, 646–651.

Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* 33, 661–667.

Wong, A.S., Choi, G.C., Cheng, A.A., Purcell, O., and Lu, T.K. (2015). Massively parallel high-order combinatorial genetics in human cells. *Nat. Biotechnol.* 33, 952–961.

Wong, A.S., Choi, G.C., Cui, C.H., Pregelbig, G., Milani, P., Adam, M., Perli, S.D., Kazer, S.W., Gaillard, A., Hermann, M., et al. (2016). Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc. Natl. Acad. Sci. U S A* 113, 2544–2549.

Zhou, P., Chan, B.K.C., Wan, Y.K., Yuen, C.T.L., Choi, G.C.G., Li, X., Tong, C.S.W., Zhong, S.S.W., Sun, J., Bao, Y., et al. (2020). A three-way combinatorial CRISPR screen for analyzing interactions among druggable targets. *Cell Rep.* 32, 108020.