# Protocol for genome-wide RNAi screening using pooled shRNA library

## Introduction

Genome-wide RNAi library screen can be performed by two ways, namely, arrayed screen and pooled selection/screen. For arrayed screen, it needs to be assisted by high throughput equipments in order to screen a large number of shRNA constructs. On the other hand, the RNAi pooled selection/screen can be done in a regular lab because the experimental procedures don't need high throughput instruments for proceeding. The aim of this protocol is to provide a fundamental consideration and guidelines for performing genome-wide RNAi pooled selection/screen.

To successfully conduct an RNAi pooled screen/selection, it is extremely important to perform the screen in a condition that most of the transduced cells survived in the selection population receive approximately one copy of shRNA or one lentivirus. Otherwise, potential candidate shRNA may co-select with other "unwanted shRNA(s)" in the same cell. This unwanted shRNA(s) will lead to serious signal noise or loss of the potential candidates during data analyses.

To establish the principle that what MOI shall be used for transduction in order to get suitable condition for the RNAi pooled screen, i.e., most of infected cells received one copy of shRNA or one virus, virologists borrowed the equation of Poisson Distribution to model or to design the experimental condition. The Poisson equation formulates as follow:

$$P(n) = \frac{m^n \cdot e^{-m}}{n!}$$

**Where**

$m$ is the multiplicity of infection or MOI (the ratio of infectious agent to cell);
$e$ is the natural exponent (approximately 2.7183);
$n$ is the occurrence of event that virus(s) enters into the cells;
$P(n)$ is the probability that a cell will get infected by $n$ viruses.

For instance, the probability that a cell got no virus or one virus under the conditions of MOI equals to 0.1, 0.2, 0.3, 0.5, and 1 is calculated according to the formula. In addition, the infected population is derived by the formula of **1-P(0)** which is also expressed as **P(infected)**, from where the probability of a cell got infected by greater than or equal to two viruses ($\geq 2$) can be derived from the formula, **P(infected)-P(1)**. (If express as percentage, the formula could be expressed as {**P(infected)-P(1)/P(infected)**}×**100%**.) Those values calculated are tabulated as follow:

| MOI | P(0) | P(1) | 1-P(0) = P(infected) | P(≥2) [% in infected population] [P(infected)-P(1)/ P(infected)] |
|-----|------|------|------|------|
| 0.05 | 0.9512 | 0.0476 | 0.0488 | 2.46 |
| 0.1 | 0.9048 | 0.0905 | 0.0952 | 5 |
| 0.2 | 0.8187 | 0.1637 | 0.1813 | 9.7 |
| 0.3 | 0.7408 | 0.2222 | 0.2592 | 14.3 |
| 0.5 | 0.6065 | 0.3033 | 0.3935 | 23 |
| 1.0 | 0.3679 | 0.3679 | 0.6321 | 42 |
| 3.0 | 0.0498 | 0.1494 | 0.8506 | 82 |

The cell numbers required for the RNAi pooled screening are reverse proportion to MOI used. Practically, MOI equals to 0.2 or 0.3 will be recommended to perform large-scale RNAi pooled screening because MOI equals to 0.1 doesn't reduce the probability of $P(n \geq 2)$ much but it needs much more cells to reach that condition.

Polybrene is a common polycation that increases the infectivity of lentivirus by 3-5 folds (unpublished observation). The cationic polymer enhanced virus adsorption and transduction by neutralizing the charge in between viral envelope and cellular membrane. However, polybrene has potential to facilitate virus aggregation which raises the possibility of multiple virus infection during pooled RNAi screen. This character of aggregation caused by polybrene increases the probability of gaining more than one lentivirus in a cell. Thus, polybrene-assisted transducing cells are not suitable for genome-wide pooled RNAi screen.

In addition to these considerations, the following experimental conditions are also important for conducting RNAi pooled screening. They are described as follows:

**(1) Optimization of puromycin selection condition**

To generate a fully transduced population of cells for analysis, it is important to determine the minimum amount of puromycin required to eliminate untransduced cells. This is accomplished by performing a puromycin kill curve to determinate the concentration of puromycin needed to eliminate the untrasduced cells. For puromycin selection, the minimum antibiotic concentration used is the lowest concentration that kills 100% of untransduced cells and maximal survival transduced cells in 48-72 hours. (Please refer to puromycin kill curve protocol)

(2) **Re-titrate viral titer in cell used for screening**

The susceptibility of various cell types to virus infection (in this case means VSV-G peudotyped lentivirus) is different. To determine the amounts of lentivirus necessary to achieve the desirable MOI on the recipient cells of interest, re-titration of the pooled lentivirus obtained from the National RNAi Platform is needed. (Please refer to viral titer estimation protocol : RIU method)

(3) **Polybrene**

The use of polybrene (hexadimethrine bromide) during lentivirus transduction is improper while performing a pooled RNAi screen as mentioned above.

The protocols described below provide the instructions on how to transduce pooled lentivirus into target cells and guidelines for the isolation of genomic DNA and the preparation of PCR products amplified from genomic DNA for deep sequencing.

**Protocols**

## I. Large-scale infection by pooled shRNA lentiviruses

Not all cellular systems can be performed by the method of the RNAi pooled screening. Cell-based systems with the natures of positive or negative selection are able to be conducted by the RNAi pooled screening. Positive selection means that the copy numbers of shRNA candidates in a particular selection condition are over the non-reactive shRNAs or control shRNAs. On the other hand, the copy numbers of shRNA candidates in the condition of negative selection will be gradually reduced or even vanished. Usually, the ID numbers of shRNAs being analyzed in the final stage of selection are still huge (almost the same as original input). But the ID numbers of shRNAs can be very few in some positive selection condition, for instance essential genes required for virus replication or genes involved in apoptosis pathway triggered by some inducers. To get meaningful results, the following protocol uses hTRC1&2 110k pooled lentiviruses (C6-10) as an example to describe. In addition, each shRNA (or corresponding shRNA expressing lentivirus) is duplicate to 250-500 folds (coverage is 250-500) to guarantee that effective shRNA will not lose during transduction and selection.

1. Re-titrate the titer of shRNA lentivirus obtained from the National RNAi Platform in the recipient cells of interest using the experimental conditions being performed in your lab, for example spin infection or not. **(It is extremely important using the MOI titrated in such experimental conditions to perform the RNAi pooled screening.)**
2. Calculate the cell numbers needed.   For example, given MOI = 0.3 and coverage = 250, then the cell numbers required are $9.17 \times 10^7$ (110,000×250/0.3).
3. Seed cells onto 150 cm$^2$ dish or flask (with a cell density of 60%-80% confluence at the second day just before transducing lentivirus). (To reach 60%-80% confluence, the numbers of the cell required are dependent on cell type used.)
4. Add lentivirus needed. (Total cell numbers multiply MOI used.) At this stage, you may perform buck culture or divide into 11 cultures or 3 sub-pools' viruses as a culture. (There are 11 sub-pools of hTRC1&2 pooled lentivirus.)
5. Incubate the infected cells at 37$^o$C for 18-24 hours.
6. Replace medium with fresh complete culture medium containing optimized concentration of puromycin. Return cells to $CO_2$ incubator for selection and change medium (containing puromycin) every three days.
7. Expand and freeze the cells established.

II. Library screen/ selection

1. Seed shRNA-expressing cells onto 10 cm$^2$ (cells from one sub-pool shRNAs) or 150 cm$^2$ (co-culture 3 sub-pools shRNA-expressing cells) dish or other experimental conditions for particular phenotype selection.

2. Subject to phenotype selection (highly depend on experimental designs).

3. If the experiment is a long-term selection, keep the cells at an optimal condition (passage the cells regularly). You may take out $2.75\times10^7$ to $5.5\times10^7$ cells ($110,000\times250$ to $110,000\times500$ for the compensation of DNA loss during purification) every week (the period is dependent on the experimental designed)

4. In the special positive selection that only a few cells survived, expand the cells to above $7.5\times10^6$ to $1.5\times10^7$ cells for the isolation of genomic DNA.

## III. Identification of hits

To identify the potential hits, the ID of shRNAs in selected cells can be determined by deep sequencing or microarray approach. In this protocol, only the method of deep sequencing is mentioned.

The genome (with shRNA sequence) of lentivirus integrates into host chromosome after transduction. Thus, shRNA sequence can be passed down to daughter cells during cell proliferation. Therefore, the whole idea of determining shRNA ID is PCR amplification of shRNA sequence integrated in genomic DNA and followed by sequencing. In general, 180 μg to 200 μg is the minimal amounts of DNA needed for the deconvolution of the ID of the shRNAs in the selected population. The amounts are calculated by the following formula: the numbers of shRNA of the library used × coverage × the mass of single cell. (For instance, when the library of 110K shRNAs is subjected to screening, the amounts of the DNA would be $110,000 \times 250 \times 6.6 \times 10^{-6} = 182$ μg.) However, if only a few cells left after selection, 20 μg to 30 μg of DNA is sufficient to decode the ID of the shRNAs selected.

The protocol described here includes: (i) isolation of genomic DNA; (ii) PCR amplification of shRNA; (iii) restriction enzyme digestion and gel purification of the PCR products.

**It is important to pay attention to note** that the PCR products prepared for deep sequencing need that both 5'-ends of PCR products contain phosphor group and 3'-end with OH group. OH group allows the PCR products can be tagged with A nucleotide by taq polymerase and phosphor group allows the PCR products can be ligated to an adaptor (adaptor needs for further PCR amplification of the PCR products for subsequent deep sequencing) by TA pairing method.

### A. Purification of genomic DNA from selected cells

1. Wash cells with ice-cold PBS twice, and then scrap the cells from dish into 15-ml centrifuge tube. For suspension cells, collect the cells from cultural flask into 15-ml centrifuge tube and wash the cells twice using step 2 condition.
2. Pellet down cells by centrifugation at 1,000-2,000 rpm for 5 min at 4°C.
3. Discard the supernatant and resuspend the cells with 3-5 ml TNE or PBS buffer, leave the tube at RT.
   {TNE buffer (50 mM Tris-HCl, pH8.0; 1mM EDTA; 150 mM NaCl; 1 mM Na-azide)}
4. Add SDS (final concentration: 0.5 %.)
5. Mix gently by inverting the tube for several times.
6. Add RNaseA (final concentration: 100-200 μg/ml).

7. Mix gently by inverting the tube for several times, and then incubate at 37°C for 30 min.

8. Add proteinase K (final concentration: 100-200 μg/ml).

9. Mix gently by inverting the tube for several times, and then incubate at 37°C for O/N or 52°C for at least 2 hrs.

10. Extract once with equal volume of phenol (pH8.0), followed by phenol/chloroform extraction till the interface is clean, finally once with chloroform. (Do NOT vortex! Instead, rotating the mixture on a rotating disc for at least 1 hr or longer.) Centrifuge at 4,000 rpm for 15-30 min for separating aqueous and organic phase. (Sorvall super T21, or equivalent ST-H750 rotor.)

11. Precipitate DNA by adding 2 volume of absolute ethanol, 1/10 volume of 3M NaOAc (pH5.2), and mix well.

12. Transfer DNA aggregates into enppendorf tube and wash DNA twice with 70% alcohol.

13. Bring down the DNA by centrifugation at 10,000 rpm for 1 min, and remove residue EtOH as much as possible.

14. Air dry.

15. Dissolve the DNA in 0.1x TE buffer (pH8.0) with appropriate volume to let **the concentration to be approximately 1 μg/μl**. (After completely dissolving, DNA should be very sticky!!)

16. Determine DNA concentration and stock at 4°C ready for use.

**B. PCR amplification and gel purification**

5 to 10 μg of PCR products are needed for deep sequencing. To minimize bias caused by PCR amplification, the protocol has been developed/optimized by single round of PCR. The size of the PCR products is 350 bp in length. The regions of amplification and restriction fragments being purified are depicted in the following diagram:



5'U6p-all/F: aggcgcg**ccga**gggcctatttcccatg/ aggcgcg**ccagaga**gggcctatttcccatg
LKO1&5/R: tgtggatgaatactgccatttgtctc

The detailed procedures are described as follows:

**I.    PCR amplification of shRNAs sequence from genomic DNA**

1.  Optimize genomic DNA amounts and PCR cycle (shall be determined for each batch of genomic DNA). Typical results are shown in Fig 1 and Fig 2 below:
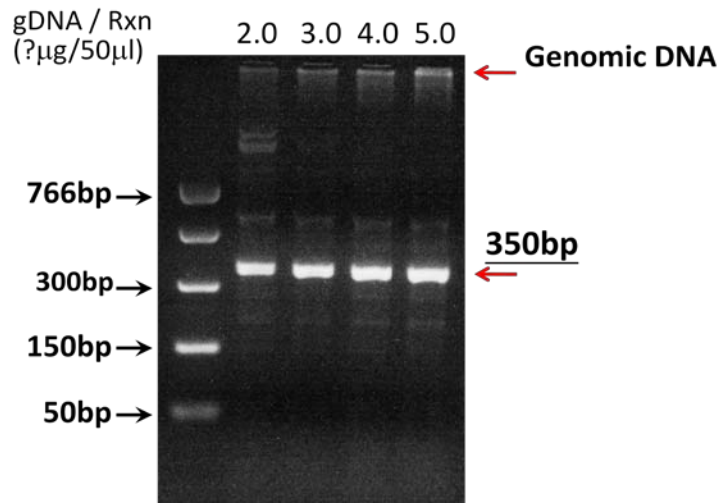


**Fig 1. <u>Amounts of genomic DNA used (data from 28 cycles).</u>** Titrate genomic DNA amounts used for PCR (per 50 μl reaction). As figure indicated, 2 to 5 μg per reaction in 50 μl is subjected to PCR amplification.
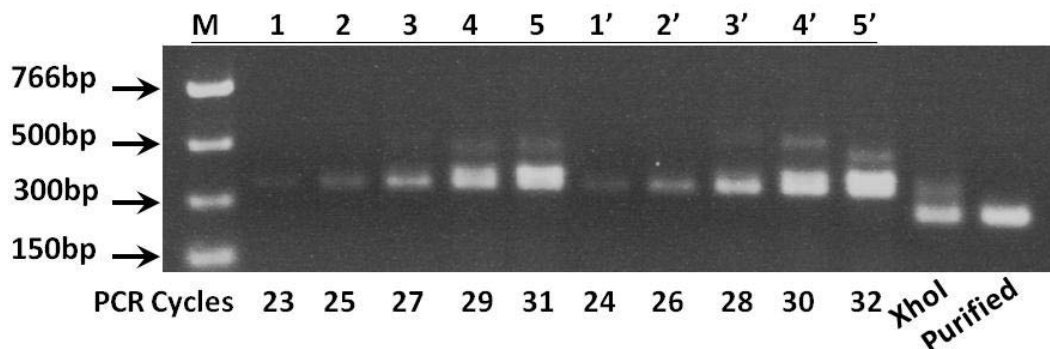


**Fig 2. <u>Optimization of PCR cycle.</u>** To minimize bias caused by over amplification of PCR, the exponential amplification was determined by a kinetic PCR. As indicated, various PCR cycles (23, 25, 27, 29, 31) were performed. The reactants were re-added primers, dNTPs and enzyme (the same as original amounts) and amplified for one more cycle (24, 26, 28, 30, 32) to enrich the perfect complimentary PCR products (as indicated by the lower

band of lane 5' and others). The upper band of the PCR products was derived from DNA annealing of different shRNA species. The resultant annealing products form bubble at the region of shRNA sequences, which lead to slower migration rate, and are resistant to XhoI digestion as indicated by the upper band of the lane of XhoI.

"Purified" indicates the DNA fragments purified from gel after AscI and XhoI digestion.

The results show that PCR products of cycle 28 are good quality for NGS in terms of the following two reasons: (i) PCR maintains at exponential stage; (ii) good products for XhoI digestion.

2. Set up PCR master mix as follows:

| Component | Vol (μL) | Vol (μL) |
|---|---|---|
| 5x KAPA HiFi Fidelity Buffer (contains 2 mM $Mg^{2+}$ at 1x) | 20 | 130 |
| dNTP Mix (10 mM each dNTP) | 2 | 13 |
| Forward Primers for TRC1 and TRC2 mix ( 100 μM) | 1 | 6.5 |
| Reverse primer (100 μM) | 1 | 6.5 |
| KAPA HiFi DNA polymerase (1U/μl) | 2 | 13 |
| gDNA Template(1 μg/μl) | 2-4 | 26-65 |
| Add dH$_2$O to | 100 | 650 |

Primers used:

5' U6p-TRC1/F: aggcgcgccgagggcctatttcccatg (target TRC1 vector)

U6p-TRC2/F: aggcgcgccagagagggcctatttcccatg (target TRC2 vector)

LKO1&5/R: tgtggatgaatactgccatttgtctc

3. Add the above reagents sequentially into eppendorf tube (**except DNA polymerase**) and mix thoroughly by vortex.

4. Add DNA polymerase after brief spin, and then gentle mix the mixture followed by spin down the mixture.

5. Aliquot 50 μl of the mix into six 200 μl PCR tubes.

6. Perform the PCR reaction using the following parameters:

|  | **Temp.** | **Time** |
|---|---|---|
| Pre-heat | 98°C | 2 min |
| Denaturation | 98°C | 15 sec |
| Annealing | 69°C | 30 sec |
| Extension | 72°C | 20 sec |
| Cycles | 28 cycles | |
| Final Extension | 72°C | 5 min |

7. Pool the contents of amplifying products together.
8. Analyze by 1-1.5% agarose gel to check the quantity and quality of the PCR products.
9. Spin column purification to remove primers and dNTPs, and to concentrate PCR products to appropriate volume with 0.1x TE buffer. The products are ready for restriction enzyme digestion.

## II. Digestion of the PCR products with XhoI and AscI restriction enzymes

1. Set up the restriction enzyme reaction to digest PCR products as follows:

   PCR products --------------------------------- 650 μl
   10x FastDigest Buffer ------------------------ 78 μl
   FastDigest XhoI (進階 Cat#: FD0695) ----- 20 μl
   FastDigest AscI (進階 Cat#: FD1894) ---- 20 μl

2. Incubate the mixture at 37°C for 5 hours to overnight.
3. Analyze 3 μl of the digested mixture using 1.5% agarose gel electrophoresis to ensure the digestion is complete.
4. Add dNTP mix (10 mM each) to final concentration of 0.2 mM followed by adding 100 units of Klenow fragment of DNA polymerase I.
5. Incubate at RT for 30 min to fill-in the protruding ends.
6. Concentrate the digestion mixtures by passing them through 6 Qiagen spin columns (QIAquick Cat. # 14208). This procedure also gets rid of unwanted small fragments that will interfere the subsequent gel purification.
7. Elute the DNA with 60 μl 0.1x TE buffer and pool all eluents together.
8. The sample is ready for gel purification.

**IV. Isolation of digested products from agarose gel**

1. Cast three sealed wells with 1.5% NuSieve 3:1 agarose (Lonza) containing ethidium bromide (0.5 μg/ml) with a volume capacity of 130 μl per sealed well (TAE buffer give rise to better resolution).
2. Separate the DNA fragments at 50 volts for above 2 hrs.
3. Locate the desired bands with a long-wave (355-360 nm) portable UV lamp.
4. Cut out the region of the desired DNA band with a razor blade. Chopped the gel slice into small pieces as much as possible and transfer the sliced gel into a 1.5 ml microfuge tube.
5. Add 1.5 to 2 volumes of **<u>Tris-saturated phenol (pH 8.0)</u>** into tube vortex for 30 seconds. Incubate the tube at RT for 10 min, then add 20 μg of glycogen (Roche). Freeze the content at -80°C for at least 5 hrs to overnight.
6. Centrifuge at 13,000 rpm in a table top centrifuge for 30 to 60 min at 18°C.
7. Transfer the upper layer of aqueous phase to a new microfuge tube.
8. Extract the aqueous solution once with equal volume of saturated phenol [pH8.0], once with phenol/chloroform, finally once with chloroform. At each step shakes the tube by vortex for 30-60 seconds vigorously and centrifuge at highest speed in a table top centrifuge for 30 min at 4°C.
9. Finally, transfer the aqueous phase to a new microfuge tube, add 0.1x volume of 3 M sodium acetate (pH 5.2) and 2 volumes of absolute ethanol and put the tube at -20℃ for overnight.
10. Centrifuge the tube at 13,000 rpm for 10 min at 4℃.
11. Wash the pellet with 70% ethanol twice, air dry and dissolve the pellet in 30 μl of 0.1x TE. (adjust the DNA concentration to be approximately 200 ng/μl.)
12. Determine DNA concentration (usually the ratio of 260/280 shall be in the range of 1.85-2.0; and 260/230 shall be greater than 2.3) and check the quality of the DNA by taking out 50-100 ng of DNA and runs the sample in 1.5% agarsoe gel.
13. The DNA is ready for deep sequencing.

**Links**

1.  Puromycin kill curve protocol
    [http://rnai.genmed.sinica.edu.tw/file/protocol/19_PuromycinKillCurve.pdf](http://rnai.genmed.sinica.edu.tw/file/protocol/19_PuromycinKillCurve.pdf)

2.  Viral titer estimation protocol: RIU method
    [http://rnai.genmed.sinica.edu.tw/file/protocol/4_1_EstimationLentivirusTiterRIUV1.pdf](http://rnai.genmed.sinica.edu.tw/file/protocol/4_1_EstimationLentivirusTiterRIUV1.pdf)

**Links**