Created by Chong Park
Adapted mainly from Kampmann et al., *Nature Protocols* 2014

Considerations for screen design

Our strategy for screens has performed very well for screens comparing the growth of untreated cells with the growth of cells in the presence of a drug or toxin that causes an intermediate selective pressure. By contrast, treatments that kill more than 50% if the cell population at given time should be avoided and instead several rounds of moderate treatment should be used. Stronger selections will be more stringent for genes that confer strong resistance to drug but will increase stochastic noise by introducing bottlenecks, and will prevent any reliable measurement of genes that confer sensitivity to drug. Drugs should typically be administered as pulsed doses to introduce a temporary selective pressure and then recovered to allow resistant cells to grow out. Pulses can be controlled by resuspending cells in fresh media or simply diluting the culture over time.

The most important factor in obtaining reliable data from pooled screens is the representation of library elements in the cell population. Aim for 1000X coverage during the screen and do sequencing 500X or lower.

To distinguish the effect that sgRNAs have on cell growth under standard (untreated) conditions from treatment-specific effects, experiments should be designed to compare the population at the time point before selection (t0), the treated population at the endpoint of the experiment and a control untreated population that grew for the same time in parallel. Inducible expression of the system to prevent losing essential genes from the pool before selection is a nice feature but can be difficult to implement as dox-inducible systems can be leaky, silenced, or toxic. Selecting cells promptly and starting the screen 5-6 days post-infection will prevent too much loss of essential gene-targeting sgRNAs.

Quantification of the frequencies of cells expressing a specific sgRNA in starting populations and in untreated and treated populations can be used to determine different sgRNA phenotypes, which we have previously defined quantitatively. In brief, gamma phenotypes quantify the effect of an sgRNA on cell growth in the absence of treatment, tau phenotypes quantify the effect of an sgRNA on cell growth in the presence of treatment and rho phenotypes quantify the effect of an sgRNA on resistance to the treatment. Although growth and selection conditions must clearly be optimized for each screen and cell type, several general principles should guide screen design:

- Selection should be set up so that a sufficient dynamic range between treated and untreated conditions can be expected. In our screens, a difference of around six net doublings between treated and untreated conditions worked

well to see major differences.

- Given the complexity of the sgRNA library, it is extremely important to avoid tight population bottlenecks at each stage of the experiment: lentiviral infection, screen selection and preparation of the sequencing sample. We aim to maintain 1,000-fold representation of each library sgRNA on average at each stage. However, it is often difficult to conduct the infection at 1,000-fold coverage (as the target infection rate is 20-50%), so we infect at around 300-fold coverage, select, and then harvest "T0" post-selection. Any noise introduced with this infection bottleneck is largely ignored as it occurs pre-T0.
- For example, for a screen of an sgRNA library with 50,000 different sgRNAs, the cell population should be maintained at 50 million cells or more. In particular, when designing a screen, care should be taken to avoid selective pressures that are so high that a large fraction of the cell population is killed. We aim for selective pressures that do not kill more than 50% of the population at any time point during the screen.
- Suspension cell lines facilitate culturing of large cell numbers, as they can easily be grown in large spinner flasks or bioreactors, and they do not require trypsinization for passaging but can simply be diluted. We had good experiences with the human K562 leukemia cell line, which does not grow in clusters and grows stably over many generations. Other suspension lines, including B cell, T cell and myeloma lines have also worked well in our hands. Adherent cell lines can be seeded on microcarriers and then grown in spinner flasks or bioreactors, although this requires some individual optimization up front for the individual cell line.
- We have successfully detected both positive and negative selection. However, it can be more challenging to confidently detect sgRNAs that cause cells to drop out of the screen, especially if they also affect growth under standard conditions, as phenotypes for such sgRNAs are then calculated on the basis of low numbers of reads in the sequencing data, which inherently will increase statistical noise.

Considerations for sample preparation

Adequate representation of the complex cell population should be maintained during library preparation. Genomic DNA is isolated from a number of cells that is 1,000-fold the number of different sgRNAs in the screened library. To reduce the amount of genomic DNA used as a PCR template, we size-fractionate the DNA by a restriction enzyme digest (with SbfI); this is followed by gel purification of the size range around the expected fragment size of the sgRNA expression cassette. In cases of small genomic DNA input quantities this step can be omitted, but test the PCR conditions on an inconsequential sample before using any precious DNA input

In the PCR-amplified material, the sgRNA-encoding region is highly enriched compared with its abundance in pre- PCR genomic DNA. Therefore, it is critical to avoid cross- contamination of pre-PCR material with any post-PCR material. Ideally, the pre-PCR steps should be performed using separate equipment (e.g., pipettes and vacuum manifolds) from that used for post-PCR steps.