Methods S1: Notes on screening throughput, choice of cell model and choice of vector backbone for optical pooled screens

Note 1: Optical screening throughput………………………………………………………...1

Note 2: Considerations when choosing a cell line for screening…………………………..2

Note 3: Choice of vector backbone for optical pooled screens….………………………...3

**Note 1**: Optical screening throughput

Imaging time plays a major role in determining the throughput of optical pooled screening. For a given library size and target cell coverage per perturbation, throughput (perturbations/unit time) is determined by the fraction of cells included in analysis, cell density, optical field of view and time spent per field of view.

The fraction of imaged cells included in analysis is determined by the cell mapping rate (fraction of cells assigned to a perturbation) and fraction of cells with called phenotypes. As we generally obtain >80% per-read mapping accuracy (Figure 2D), the leading cause of cell dropout in screens is a lack of mapped reads. The read count distribution is determined by the cellular expression of barcodes as well as the padlock detection efficiency. We found the EF1a promoter yields a high fraction of cells with mapped reads in both the lentiGuide-BC and CROP-seq vectors, across multiple cell lines (Table S6). The fraction of cells considered for analysis may be further limited by constraints on phenotypic features (e.g., we discarded mitotic cells and cells with insufficient reporter expression for the p65-mNeonGreen translocation assay).

The optimal cell density depends strongly on the cell type and phenotypic assay under consideration. For example, the HeLa cells in the screens reported here were seeded at a density of 50,000 cells/cm2 so that they reach a density of ~175,000 cells/cm2 (~90% confluency) at the time of imaging. Larger cells commonly used for imaging, such as U2-OS osteosarcoma cells, may need to be screened at reduced density.

Once barcode expression, barcode detection, and cell density have been optimized, throughput is controlled by the acquisition speed of the microscope. In this study, barcode imaging was performed on a Nikon Ti-E widefield microscope equipped with a motorized stage, single sCMOS camera and motorized filter turret (STAR Methods, Key Resources Table). Hardware was controlled using open-source microscopy software (Micromanager). At each stage position, the filter turret was rotated to image each of 5 different channels (four sequencing channels for Illumina MiSeq dyes and a DAPI nuclear stain for registration between cycles), requiring ~5 seconds per stage position. Imaging a full 6-well plate at 10X magnification required 3,198 positions to be acquired, requiring a total of 6.5 hours per sequencing cycle.

Imaging throughput may be boosted substantially by (a) using faster hardware (reduced filter-switching time, increased stage speed and reduced communication overhead), (b) increasing the imaging field of view (c) eliminating mechanical filter switching (e.g., by using multi-band filter sets and fast-switching LED/laser excitation) and (d) reducing the number of fluorescent channels for sequencing. For example, using a Nikon Ti2 microscope (2X increased field-of-view over Nikon Ti-E) running with NIS Elements software (fast communication with hardware components) addresses (a) and (b) above to increase imaging throughput 5-fold. Furthermore, we project that combining the Ti2 improvements with two-color sequencing-by-synthesis chemistry (e.g, Illumina NextSeq/MiniSeq) and a multiband filter set in conjunction with an LED light engine could increase the imaging throughput by ~10-fold. With these improvements, the imaging time per sequencing cycle could be reduced to ~30 minutes per 6-well plate.

**Note 2**: Considerations when choosing a cell line for screening

When choosing a cell line model for an optical pooled screen, a number of factors must be considered, including the fraction of cells with accurately mapped perturbations, Cas9 efficiency (for CRISPR screens) and suitability for modeling the biological process of interest. We have already demonstrated *in situ* detection in a panel of cancer cells, including adherent and suspension lines. Additionally, the maximum feasible cell density is an important consideration as it directly impacts screening throughput (i.e., number of cells imaged per unit time and cost per cell).

The perturbation mapping rate, which depends on the number of spots per cell and their brightness, is affected by barcode expression as well as *in situ* amplification efficiency. Expression level may be tuned by varying promoter choice or selection strategy (Supplemental Note 3) while *in situ* amplification is likely affected by biophysical properties of the cell after fixation (e.g., molecular constituents, organization and dynamics). Barcode amplification may be optimized by altering the *in situ* processing (enzyme and dNTP concentrations, buffer components or fixation conditions). The protocol in this manuscript achieves high yield from a number of cell lines and is a recommended starting point.

For CRISPR knockout and CRISPRi/a screens, there are well-described assays to validate Cas9 performance in the target cell line (Horlbeck et al., 2016; Joung et al., 2017). Screens may be performed in a cell line with heterogeneous Cas9 expression (e.g., a cell line produced by lentiviral transduction and antibiotic selection); however, flow-enrichment or clonal isolation (as used in this work) can boost CRISPR efficiency and improve statistical power. For CRISPR knockout screens, the frameshift reporter screen outlined in the main text is a convenient all-in-one test for Cas9 efficiency and perturbation readout in a target cell line/clone; analogous fluorescent reporter screens could be used to validate CRISPRi/a screening.

**Note 3**: Choice of vector backbone for optical pooled screens

In the main text, we describe two different strategies for CRISPR-based optical screens. The original LentiGuide-BC strategy used an sgRNA and associated barcode separated by a long (~2 kb) intervening sequence. We and others have noted that such designs can result in a substantial degree of barcode swapping that diminishes the statistical power of screens [(Feldman et al., 2018; Hill et al., 2018; Sack et al., 2016; Xie et al., 2018)](https://www.zotero.org/google-docs/?zaBab4), likely due to homologous recombination during lentiviral integration. We addressed this problem by co-packaging the plasmid library with a non-homologous carrier plasmid (pR\_LG), which eliminated barcode swaps but also reduced the functional viral titer. As a second strategy, we used the CROP-seq approach in which the sgRNA is duplicated onto a Pol II transcript that can serve as a template for reverse transcription and *in situ* detection. In this approach, lentiviral recombination is not an issue and thus the effective viral titers can be much higher (~100X).

After we optimized the oligos used for *in situ* amplification of the CROP-seq vector (Figure S2, Table S3), both strategies produced *in situ* sequencing spots that could readily be detected at 10X magnification and yielded similar screening throughput (Table S6). In addition to the improvement in titer, the CROP-seq backbone requires only one cloning step to generate sgRNA libraries (versus two steps for LentiGuide-BC). Although the CROPseq screens in this manuscript use an earlier version of the sgRNA scaffold sequence, we have recently optimized padlock detection probes for a modified scaffold (CROPseq-puro-v2 in the Key Resources Table) [(Dang et al., 2015)](https://www.zotero.org/google-docs/?P2Pd6b) that has been shown to improve CRISPR efficiency (Table S3).

For these reasons, we generally recommend the CROP-seq backbone for most CRISPR-based optical screens. For applications where an associated barcode is advantageous or necessary, such as barcoded ORF libraries or dual-sgRNA vector backbones, the LentiGuide-BC flanking sequences may be useful for *in situ* detection.

Promoter choice and selection strategy also affect the abundance of expressed barcodes. In our initial screens, we used a CMV promoter that resulted in a broad distribution of barcode transcripts in HeLa cells. Comparatively, an EF1a promoter yielded a more uniform distribution in HeLa cells and is a strong promoter in many cell types. The expression of barcodes can be further boosted by using a more stringent antibiotic selection such as puromycin or zeocin [(Nakatake et al., 2013)](https://www.zotero.org/google-docs/?fM8jBK) or by FACS-sorting a co-expressed fluorescent marker (Figure S2).