Supplementary Note 1. **Selection of capture sequences for direct capture Perturb-seq and construction of sgRNA expression plasmids.**

To design capture sequences, we considered three important criteria: (1) the presence of these sequences in sgRNAs (to allow for primer annealing during RT) should not compromise CRISPR activity, (2) these sequences should facilitate robust and specific barcoding of targeted RNAs, and (3) addition of target-specific primers carrying the reverse complement of these sequences to single-cell reactions should not interfere with the generation of gene expression profiles. To identify promising capture sequences for our 3' scRNA-seq platform, we used CRISPRi in K562 cells to screen candidate sequences. We assessed effects on guide activity conferred by incorporation of candidate sequences into one of two positions in guide constant regions: in stem loop 2 or at the 3' end (**Supplementary Fig. 1c,d**). In agreement with previous observations[1], we found that extra sequence in the loop region of the CR1 constant region[2,3] minimally impacted guide activity, while incorporation of extra sequence near the 3' end often decreased activity (**Supplementary Fig. 1d**). We chose two capture sequences, capture sequence 1 (cs1) and capture sequence 2 (cs2), to evaluate for direct capture Perturb-seq because incorporation of either cs1 into the loop of CR1 ($CR1^{cs1}$) and cs2 at the 3' end of CR1 ($CR1^{cs2}$) does not compromise guide activity (**Supplementary Fig. 1c,e**). We refer to guides carrying these constant regions as sgRNA-$CR1^{cs1}$ and sgRNA-$CR1^{cs2}$. Notably, we found that while the position of cs2 is interchangeable, incorporating cs1 at the 3' end decreases guide activity (**Supplementary Fig. 1f**). *Therefore, we strongly recommend against the use of sgRNAs with cs1 at the 3' end.* Of note, guides carrying our optimized capture sequences can also be used for 5' direct capture. To demonstrate this, we performed 5' direct capture with sgRNA-$CR1^{cs1}$ using a guide-specific RT primer (oJR161) which contains the reverse compliment of cs1 (**Fig. 1c,d** and **Supplementary Fig. 2a,c,d,h, 3a**).

In another iteration of 5' direct capture Perturb-seq (**Fig. 1b, Supplementary Note 4**), we designed a guide-specific RT primer (oJR160) to target a sequence present within standard guide constant regions. Notably, we designed this primer to bind across many commonly-used guide variants, making our approach amenable to screening across *Streptococcus pyogenes* Cas9 (spCas9) perturbation systems, including those with modified guide libraries such as SAM-based CRISPR activation (CRISPRa)[1]. In principle, this approach is also readily adaptable for use with guides from other species and for other classes of Cas proteins.

We constructed many sgRNA expression plasmids to evaluate the effects of incorporating different capture sequences into the constant regions of sgRNAs (**Supplementary Fig. 1d**, **Supplementary Table 1**). Among these are vectors pAX064-pAX099, pBA896, pBA899-pBA904, and pBA970, which encode GFP-targeting sgRNAs (EGFP-NT2 targeting region, 5'-GACCAGGATGGGCACCACCC-3')[4] with variants of the constant region sequence CR1. As described above, these variants contain capture sequences at the 3' end (prior to the terminating poly-T tract) or within the loop region of the so-called “stem loop 2” (**Supplementary Fig. 1c**). Loop modified constant regions also contain an extension to the stem region as found in sgRNA 2.0[1]. To construct these vectors, we replaced the sgRNA constant region sequence in pU6-sgRNA EF1Alpha-puro-T2A-BFP (Addgene, #60955) with our modified variants. More specifically, to make pAX064-pAX099 (**Supplementary Table 1**), we inserted synthesized CR1 variants by Gibson assembly using the BstXI and XhoI sites of the parental vector. For pBA896, pBA899-pBA904 and pBA970 (**Supplementary Table 1**), synthesized constant region variants were inserted using BlpI and XhoI. Vectors pBA900 and pBA904 contain sgRNA-$CR1^{cs2}$ and sgRNA-$CR1^{cs1}$, respectively.

To construct our dual-guide expression vectors, we made and tested an additional six sgRNA constant region variants using published constant region sequences CR2 and CR3[2] (**Supplementary Fig. 5a)**. These are encoded in pJR73-pJR78 (cloned as described above for pAX064-pAX099). Using GFP depletion as a readout of sgRNA activity, we determined that all of these new variants maintained high CRISPRi activity (**Supplementary Fig. 5a**). For downstream dual-guide Perturb-seq experiments, we constructed guides with two variant constant regions, CR3 with cs1 in stem loop 2 ($CR3^{cs1}$) and CR2 with cs2 at the 3' end ($CR2^{cs2}$), and paired them with sgRNA-$CR1^{cs1}$ in dual-guide vectors. This design, using distinct constant region variants in positions A and B, was motivated by previously published multi-guide vector system[2] and serves to minimize intramolecular recombination between guide sequences during lentiviral transduction. We refer to these dual guide vector configurations as $CR3^{cs1}$/ $CR1^{cs1}$ and $CR2^{cs2}$/ $CR1^{cs1}$.

To test capture of our dual-guide vector expressed sgRNAs, we performed direct capture Perturb-seq with our dual-guide vectors. At a constant sequencing depth, we found that sgRNA-$CR3^{cs1}$ produced

10-fold higher index capture than sgRNA-CR2^{cs2} (sgRNA-CR3^{cs1} median of 776 UMIs/cell; sgRNA-CR2^{cs2} median of 74 UMIs/cell) (**Supplementary Fig. 5b**). These results are consistent with our observations from single-guide sgRNA-CR1^{cs2} vectors (**Supplementary Fig. 2a**) that 3' guide modification is more destabilizing than modification of the loop.

Finally, we also constructed two CROP-seq vectors to use as controls in our GFP-depletion assay (**Supplementary Fig. 1d**). These are pBA950 (labeled as "CROP-seq modified for CRISPRi") and pBA960 (labeled as "CROP-seq"). These vectors were derived from CROPseq-Guide-Puro (Addgene, #86708) as follows: First, an intermediate vector (pBA948) was constructed by replacing the existing selectable marker in CROPseq-Guide-Puro with BFP (synthesized dsDNA inserted by Gibson assembly using PstI and MluI). Then, pBA960 was made by programming the encoded sgRNA with EGFP-NT2, and pBA950 was made by replacing the entire human U6-driven sgRNA expression cassette with the sgRNA expression cassette from pU6-sgRNA EF1Alpha-puro-T2A-BFP (synthesized dsDNA inserted between PpuMI and NcoI by Gibson assembly). This cassette is driven by a modified mouse U6 promoter and encodes a BlpI-containing sgRNA(F+E) programmed with EGFP-NT2. To allow cloning of pBA950, a 21 bp repeat outside of the lentiviral LTRs was also removed.

Supplementary Note 2. **Cloning single-guide direct capture Perturb-seq libraries.**

For our single-guide direct capture Perturb-seq experiments, we made four CRISPRi libraries: the UPR GBC, UPR sgRNA-CR1^{cs1}, UPR sgRNA-CR1^{cs2}, and iPSC sgRNA-CR1^{cs1} libraries (**Supplementary Tables 2, 3**). Each of these was constructed with one of three guide constant regions: our parental constant region CR1 (as in pU6-sgRNA EF1Alpha-puro-T2A-BFP), CR1 with cs1 inserted into stem loop 2, or cs2 appended to the 3' end (**Supplementary Note 1**). We made these libraries by cloning target seqeunce-containing inserts (annealed oligos from IDT with BstXI/BlpI overhangs) into pBA571, which is the Perturb-seq GBC library[2], or one of two expression vectors, pBA900 (CR1^{cs2}) or pBA904 (CR1^{cs1}) (**Supplementary Note 1**) after digestion with BstXI and BlpI (https://weissmanlab.ucsf.edu/links/sgRNACloningProtocol.pdf). Library vectors were then clonally isolated and verified by Sanger sequencing of the protospacer region and, for pBA571-derived vectors, the corresponding GBC. Notably, pBA571-derived vectors encode guide expression cassettes in the opposite orientation to those in pBA900 and pBA904.

Supplementary Note 3. Protocol for 5' direct capture Perturb-seq of unmodified sgRNAs.

5' direct capture Perturb-seq platform is suitable for use with standard sgRNA sequences (no engineered capture sequence). Although oJR160 should anneal to most spCas9 perturbation systems, a wide variety of sgRNA variants are currently in use, so it is necessary to check oJR160 compatibility before proceeding. Additionally, if running multiple lanes on the 10x Genomics Chromium Controller, oJR165 variants with unique i7 indices will be necessary to demultiplex libraries during sequencing.

This protocol follows the Chromium Single Cell V(D)J Reagent Kits User Guide (10x Genomics, CG000086). The Chromium Single Cell 3' Reagent Kits v3 User Guide with Feature Barcoding technology for CRISPR screening (10x Genomics, CG000184) serves as a reference for the cDNA SPRI cleanup.

Primer designs for 5' direct capture Perturb-seq:

Primer ID	Primer sequence (5' to 3')
oJR160 (RT primer)	Design: 10X sequence–guide annealing AAGCAGTGGTATCAACGCAGAGTACCAAGTTGATAACGGACTAGCC
oJR163 (Amplification primer)	Design: P5–TruSeq Read 1 / Index 2 (on bead sequence) AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
oJR165 (Amplification primer)	Design: P7–**i7 index**–Nextera Read 2/Index 1–Shared Nextera Read 2/Annealing–annealing region CAAGCAGAAGACGGCATACGAGAT**AGGAGTCC**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGTACCAAGTTGATAACGGACTAGCC

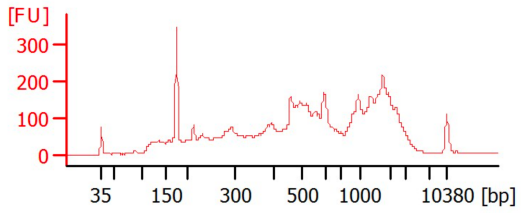
1) **Perform droplet capture with the 10x Genomics Chromium Single Cell V(D)J Reagent Kits.** Prior to droplet formation, we add oJR160 (guide-specific RT primer) to the RT Master Mix**.** We add 5 pmols oJR160 (0.5 uL of 10 uM) directly to the 68.3 uL of RT Master Mix used for each 10x lane.

2) **cDNA amplification**. oJR160 has an adapter identical to the adapter sequence on the Poly-dT RT Primer (10x Genomics, PN-2000007). This adapter servers as a primer binding site for the Non-Poly(dT) primer (10x Genomics, PN-220106) during cDNA amplification, and thus allows amplification of reverse transcribed guides to occurs concurrently with standard cDNA amplification.

3) **cDNA cleanup and separation of cDNA libraries fractions.** For reference, see **Supplementary Fig. 1g** (schematic overview) and step 2.3 of our 3' direct capture Perturb-seq protocol (10x Genomics, CG000184) which generally mirrors the 5' direct capture workflow. Briefly, guide-containing cDNA amplicons (168 bp) can be enriched by size selection. To do this, we perform a 0.6X left-sided SPRI cleanup reaction and collect both the beads (which carry material for preparing the gene expression library below in step 7) and the supernatant (which is enriched for guide-derived cDNA amplicons). Use the supernatant, we complete a 0.6X-1.2X double-sided SPRI selection in order to collect the guide-containing fraction and proceed to step 4.

4) **Quantify the guide-enriched cDNA fraction.** The expected size of the guide-containing cDNA amplicons (with adapters, CBC, UMI) is 168 bp. In K562 cells, a peak is apparent in traces produced by a Bioanalyzer High Sensitivity DNA Analysis (Agilent) (representative plot below). To determine how much template to add to the specific amplification PCR in the next step, we quantify the mass of 145 bp to 1000 bp products in the sample.

Sample Bioanalyzer trace, with peak at 168 bp:



5) **Specific amplification of guide-containing cDNA amplicons.** Using 5 ng of the guide-enriched cDNA fraction per reaction, we perform 4 PCRs with KAPA HiFi ReadyMix and 0.6 uM of primers **oJR163** and **oJR165**, as below. We purify the resulting guide (index) sequencing libraries with a 0.8X SPRI cleanup reaction to remove primers and elute in 40 uL.

Components	Volume for 1 reaction (uL)	Volume for 4.25 reactions (uL)
10 uM P5-primer	3	12.75
10 uM P7-primer	3	12.75
Template DNA (5 ng)	depends on template concentration	depends on template concentration
2X Kapa HiFi Master Mix	25	106.25
Water	depends on template concentration	depends on template concentration
Total	50	212.5

Step	Time	Temperature	Cycles
Initial Denaturation	3 minutes	95°C	1
Denature	15 seconds	98°C	12X
Anneal/Elongate	10 seconds	70°C	
Final Elongation	1 minutes	72°C	1
Hold	hold	4°C	1

6) **Quantification of the final guide (index) sequencing libraries.** We measure the concentration of the final guide sequencing libraries using a Bioanalyzer High Sensitivity DNA Analysis (Agilent) at a 1:10 dilution. The expected size of the peak is 250 bp.

7) **Prepare the corresponding gene expression library.** Gene expression libraries are prepared from material collected in step 3. For this, refer to the Chromium Single Cell V(D)J Reagent Kits User Guide (10x Genomics, CG000086).

8) **Sequencing.** We sequence the guide (index) sequencing libraries as ~10% spike-ins added to gene expression libraries. The Illumina Read 2 will sequence the sgRNA constant region before sequencing the sgRNA targeting region. Therefore, it is always necessary to sequence the sgRNA library alongside a higher diversity library (such as the gene expression library) or with PhiX. We recommend using a Read 2 of 98 bp to ensure that the read sequences through the sgRNA targeting region.

Supplementary Note 4. **Cloning dual-guide direct capture Perturb-seq libraries.**

Dual-guide direct capture libraries were constructed according to the schematic depicted in **Fig. 2a**. Library components included the parental vector pJR85 (pBA904 with two BsmBI sites removed), a synthesized oligonucleotide insert (either CR3^{cs1}-hU6 or CR2^{cs2}-hU6), and a dual-guide RNA targeting sequence oligonucleotide pool (including verified PCR adapter sequences available at https://weissmanlab.ucsf.edu/CRISPR/CRISPR.html). Library oligo pools were ordered from Twist Biosciences with the following overall structure:

CR3^{cs1}/ CR1^{cs1} library: 5'–PCR adapter–CCACCTTGTTG–protospacer A–GTTTCAGAGCGAGACGTGTTTGATCTCGGGCCGTCTCAGAAACATG–protospacer B–GTTTAAGAGCTAAGCTG–PCR adapter–3'

CR2^{cs2}/ CR1^{cs1} library: 5'–PCR adapter–CCACCTTGTTG–protospacer A–GTTTGAGAGCGAGACGTGTTTGATCTCGGGCCGTCTCAGAAACATG–protospacer B–GTTTAAGAGCTAAGCTG–PCR adapter–3'

Briefly, dual-guide libraries were constructed by PCR-amplifying oligo pools. The resulting amplicons were BstXI/BlpI digested, gel extracted, and ligated into a similarly digested pJR85. The ligation products were then transformed into bacteria and amplified at scale to generate intermediate libraries (>100 bacterial colonies per library element). In parallel, either the CR3^{cs1}-hU6 insert (pJR89) or the CR2^{cs2}-hU6 insert (pJR88) were ligated into the intermediate library after BsmBI-digestion. The final dual-guide libraries were verified via NGS sequencing.

Supplementary Note 5. Design considerations for hybridization probes used to enrich targeted gene expression panels from scRNA-seq libraries.

To enrich select molecules from single-cell gene expression libraries, we used a hybridization-based protocol with carefully designed probes. These probes bind to cDNA sequences present in post-fragmentation, amplified sequencing libraries, which contain relatively small fragments derived from full-length cDNAs (~300 bp). We performed these enrichments on gene expression libraries prepared according to the Chromium Single Cell 3' Reagent Kits v3 User Guide (10x Genomics, CG000184). Although 3' scRNA-seq primarily captures the 3' end of transcripts, reads at other locations throughout the transcript are also common. Therefore, rather than rely on the assumption that reads come from annotated 3' ends, we took a data-driven approach and, using the 978 landmark L1000 genes[5], designed probes using the following procedure:

1) For each gene, we determine all transcript isoforms that account for >80% of reads in a K562 RNA-seq dataset (https://www.encodeproject.org/files/ENCFF717EVE/).

2) For each of these transcripts, we then perform a peak finding procedure to target the most common cDNA-represented sequences in cell-type matched 3' scRNA-seq data with probes. We align all scRNA-seq reads within the transcript and smooth the read density using a median filter. We then find a region that contains >80% of reads, pad the selected region with 25 bp on the 5' end and 200 bp on the 3' end, and extract this sequence to target with probes. For transcripts where the resulting sequence is >2 kb (e.g. if there is an extraneous peak of density early in a transcript), we threshold the sequence on the 5' end to make a 2 kb peak. For transcripts with insufficient sequencing coverage for empirical design, we simply target 300 bp starting at the annotated 3' end.

3) Next, for each gene, we compare the regions chosen across different transcripts. If one region is a strict subset of another, then we eliminate the smaller region.

4) Lastly, we use Twist Biosciences to design and synthesize 120 bp probes against the targeted sequence.

Supplementary Note 6. **Details of data analysis.**

Low UMI count / inviable cell removal. Droplet-based scRNA-seq data often contain small subpopulations of cells with low UMI counts, likely representing droplets that underwent inefficient reverse transcription or encapsulated only fragmented cellular debris. In our UPR experiment with 5' capture of standard sgRNAs, there was an obvious subpopulation of ~150 (~2%) cells with low UMI counts. These cells were removed by manually thresholding the number of UMIs per CBC. For analysis in **Figure 2g**, we also removed apoptotic cells. For this, we used a random forest regressor that was trained to recognize inviable cells using published data[2].

Hierarchical clustering of UPR perturbations in **Figure 1e** and cophenetic correlation. For **Figure 1e**, we generated pseudo-bulk RNA-seq profiles for each guide evaluated on each of three Perturb-seq platforms (see *Methods*). We then calculated guide-guide Spearman's rank correlation matrices for each of those datasets. We optimally ordered clusters to minimize the distance between successive leaves. To enable visual analysis, we generated heatmaps for each correlation matrix. Genes are ordered according to hierarchical clustering of the GBC data (Ward's method). To quantitatively compare the similarities between correlation matrices, we calculated a cophenetic correlation (the Pearson correlation of all pairwise similarities between perturbation profiles across datasets).

Correlation of average expression profiles in **Supplementary Figure 3b,c**. **Supplementary Figure 3b** presents Pearson correlations of pseudo-bulk RNA-seq profiles (see *Methods*) generated on different Perturb-seq platforms with guides containing the same targeting sequence. For this analysis, we calculated average expression profiles using only the top 100 most differentially expressed genes per perturbation, which we determined using random forest classifiers. Specifically, for each guide, we trained a random forest classifier (scikit-learn extremely randomized trees with 1000 trees in the forest) to predict the perturbation status of a cell, using as features the normalized expression profile (see *Methods*) of each cell for genes with a mean expression >0.5 UMI per cell. For each guide, the top 100 genes whose expression level could be used to separate perturbed and unperturbed cells in GBC Perturb-seq were considered differentially expressed. The advantage of this approach is that we do not employ a strict cutoff and therefore can assess the similarity of expression profiles across platforms regardless of the strength of the perturbation. However, this strategy always returns 100 genes per perturbation without regard for statistical significance, and thus for perturbations without robust transcriptional phenotypes, it may return genes that are not truly differentially expressed. This may deflate the apparent average correlation between platforms. To explore this possibility, we asked how the number of differentially expressed genes, as determined by a second procedure, related to the degree of correlation between expression profiles (**Supplementary Fig. 3c)**. Specifically, for each perturbation we performed a two-sample Kolmogorov–Smirnov test per gene and considered genes differentially expressed when the Benjamini/Yekutieli FDR-corrected p<0.01 for GBC Perturb-seq. As expected, we found that the number of differentially expressed genes was associated with the degree of correlation (Spearman's rank correlation: ρ=0.84, p=2e-9 for 3' sgRNA-CR1^{cs1} capture; ρ=0.86, p=5e-10 for 5' sgRNA capture).

Delineation of genetic regulons in **Figure 1f**. Previously, we used Perturb-seq to delineate genes regulated by the three signaling branches of the UPR, which are controlled by IRE1α, PERK, and ATF6[5]. In the UPR Perturb-seq experiments presented herein, we expected these same genes to be differentially expressed and to meaningfully covary based on differential UPR branch activation. We therefore used these genes to test the ability of direct capture Perturb-seq to cluster similar genetic regulons. To this end, we used the normalized expression profiles (see *Methods*) of all cells from our new UPR experiments to calculate gene-gene Spearman's rank correlation matrices for this gene set. We then quantitatively assessed the similarity of the clustering produced by calculating the cophenetic correlation, the Pearson correlation of all pairwise similarities between genes, across platforms (GBC and both 3' and 5' direct capture Perturb-seq). To generate a visual representation of the data, we next applied Ward's method to the coexpression matrix to hierarchically cluster genes, where the dendrograms tend to place coregulated genes near one another (**Fig. 1f**). Clustering revealed genes grouped by branch regulation, in agreement with our previous work.

Single-cell analysis in **Figure 1g** and **Supplementary Figure 3d**. To determine the single-cell performance of direct capture Perturb-seq, we again employed random forest classifiers. Specifically, for each

perturbation, we split our data into training (80%) and testing (20%) sets and, to separate perturbed and unperturbed (NegCtrl3) cells, trained a random forest classifier on normalized expression profiles (see *Methods*). We then tested the accuracy (balanced for perturbed and unperturbed cells) of our classifiers on the remaining 20% of cells. In this regime, the accuracies of our random forest classifiers can serve as proxies for the single-cell performance of Perturb-seq across platforms, because for strong genetic perturbations, we generally expect perturbed cells to be transcriptionally distinguishable from unperturbed cells. Using a Wilcoxon two-sided signed-rank test, we then tested for differences in classification accuracy across platforms and failed to reject the null hypothesis that direct capture Perturb-seq performs comparably to GBC Perturb-seq (Wilcoxon two-sided signed-rank test: p=0.2 for 3' sgRNA-CR1^{cs1} capture; p=0.6 for 5' sgRNA capture) (**Supplementary Figure 3d**). Next, to visually inspect single cell performance, we selected a subset of cells bearing strongly perturbative guides, specifically those targeting SEC61A1, SEC61G, ATP5B, MRPL39, CARS, HARS, and sgNegCtrl3. We reduced dimensionality by computing 10 independent components (using FastICA) followed by t-distributed stochastic neighbor embedding (t-sne) to project cells into two dimensions (**Fig. 1g**).

Cell cycle analysis in **Figure 2d**,**f** and **Supplementary Figure 5d**. To determine the most likely cell cycle phase of each cell, we used scores derived from panels of cell cycle markers as previously described[2]. This allowed us to calculate the fraction of cells in each cell cycle phase and the relative (ie. z-scored with respect to control cells) occupancy of cells in each stage of the cell cycle for subpopulations of cells.

Linear epistasis model and gene expression heatmaps in **Figure 2e,h**. We sought to determine the extent to which each dual target gene phenotype (**ab**) could be explained by independent repression of each single gene (**a** and **b**). For this, we used a previously described approach[6]. Specifically, we decomposed the transcriptional changes caused by dual perturbations into components explained by each of single perturbations by fitting a robust Theil-Sen regressor (scikit learn TheilSenRegressor with parameters fit_intercept=False, max_subpopulation=1e5, max_iter=1000, random_state=1000): $d\mathbf{ab} = c_1 d\mathbf{a} + c_2 d\mathbf{b} + e$**.** This model outputs a predicted dual-perturbation phenotype, which can then be compared to the observed phenotype. This allows detection of neomorphic behavior (i.e., gene expression changes not explained by the phenotype of either single-gene perturbation). To visualize phenotypes from our epistasis experiments, as in **Figure 2e** and **2h**, we generated a gene expression heatmap representing 50 genes. These were selected in the following manner: First, we determined genes that were differentially expressed across cells carrying the double (**ab**) and single (**a** or **b**) perturbations using a random forest classifier as previously described[6]. All genes with a mean expression greater than 0.5 UMI per cell were considered. Then, we plotted the 50 genes with the highest predictive power.

FDPS/*HUS1* single cell analysis in **Figure 2h**. To visualize the gene expression changes caused by repression of *FDPS* and *HUS1* at the single-cell level, we used uniform manifold approximation and projection (UMAP). To do this, we first determined genes that were differentially expressed across cells carrying the *HUS1/FDPS* double and *HUS1* or *FDPS* single perturbations using a random forest classifier. All genes with a mean expression greater than 0.1 UMI per cell were considered, and the top 200 genes were used to make the projection (parameters: metric='euclidean', n_neighbors=10, random_state=100, use_pca=True).

Multiplexed sgRNA analysis in **Supplementary Figure 6**. Multiplexing two sgRNAs into a single expression vector has the potential to improve CRISPRi and CRISPRa libraries in two ways: (1) higher probability of delivering at least one active guide per cell, and (2) guide synergy. To evaluate the latter, we compared the activity of distinct guides targeting the same gene (expressed from our dual guide vector) to the activity of the most active single guide in each pair (expressed from our dual guide vector alongside a negative control sgRNA) (**Supplementary Fig. 6c,d**). To rule out effects of steric competition between guides, which may occur when guides target nearby genomic sequences, we identified guide pairs in our library that target sequences <80bp apart in the genome and tested one guide from each of those pairs with an additional guide (sgRNA 3) whose genomic target is >80 bp away (**Supplementary Fig. 6e,f**).

Analysis of gene expression skew and biased in **Supplementary Figure 7a,b**. Gene expression is skewed and biased by function. To demonstrate the former, we plotted a cumulative density function of gene expression (**Supplementary Fig. 7a**). We generated this plot using scRNA-seq data from K562 cells as

follows: First, we ordered genes based on their expression level (gene rank). Then, we calculated the total fraction of mRNA UMIs accounted for by genes with a rank less than or equal to a given value. To assess bias by function, we used gene ontology (GO). For this, we first downloaded the GO SLIM dataset from Ensembl BioMart (genome build GRCh38.p13; accession date 10/08/2019). We manually curated a set of 23 GO terms of moderate size and biological interest, and then for each GO term, we calculated the fraction of the genes assigned to that term in each of 10 gene expression bins (determined by sorting all expressed genes) (**Supplementary Fig. 7b**).

Analysis of target enrichment performance **Figure 3c-e** and **Supplementary Figure 7c-h**. **Figure 3c,d** and **Supplementary Figure 7c-f**, depict analysis of scRNA-seq data generated with and without sequencing library enrichment (target enrichment) using material from our genetic interaction experiment (1 lane) and hybridization baits against the L1000 genes. To collect this data, we deeply sequenced both the unenriched, original library (3,517,544,917 reads; sequencing saturation 48%; 252,516 mean reads per cell) and the target enriched library (334,745,797 reads; sequencing saturation 62%; 26,445 mean reads per cell). The target enriched library was intentionally sequenced to only ~0.1x the depth of the original library because the L1000 genes make up 6% of the entire transcriptome. Both sequencing libraries were aligned via Cell Ranger and filtered to contain only cells bearing defined sgRNA pairs (n=6349 cells).

Figure 3e and **Supplementary Figure 7g,h** depict analysis of scRNA-seq data (n=211,103 cells before filtering for guide-identified cells) generated with and without target enrichment using material from our multiplexed CRISPRi experiment (16 lanes) and hybridization baits against the L1000 genes (unenriched 5,315,226,464 reads; enriched 3,255,486,907 reads). After Cell Ranger alignment, we filtered for cells bearing defined sgRNA pairs and identified those that produce significant transcriptional phenotypes by performing a two-sided, two-sample Kolmogorov–Smirnov test for each gene, comparing perturbed and unperturbed cells. We considered genes differentially expressed when the Benjamini/Yekutieli FDR-corrected p<0.001. For perturbations with greater than 10 differentially expressed genes, we calculated average pseudo-bulk RNA-seq profiles (see *Methods*) and then calculated a perturbation-perturbation Spearman's rank correlation matrix. We applied HBDSCAN to cluster perturbations based on the whole transcriptome data (parameters: metric='correlation', min_cluster_size=6, min_samples=1, cluster_selection_method='eom', alpha=1), and optimally ordered hierarchical clusters to minimize the distance between successive leaves (**Fig. 3e**). Additionally, to quantitatively compare the similarity of these correlation matrices, we calculated a cophenetic correlation, the Pearson correlation of all pairwise similarities between the unenriched and enriched datasets (visualized in **Supplementary Fig. 7h**).

Supplementary Note References

1. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583 (2015).
2. Adamson, B. *et al.* A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167,** 1867-1882.e21 (2016).
3. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–61 (2014).
4. Gilbert, L. A. *et al.* CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154,** 442–51 (2013).
5. Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171,** 1437-1452.e17 (2017).
6. Norman, T. M. *et al.* Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. *Science* **365,** 786–793 (2019).