Geometric coherence of single-cell CRISPR perturbations reveals regulatory architecture and predicts cellular stress

Prashant C. Raju
rajuprashant@gmail.com

Significance Statement

CRISPR screens quantify how far perturbed cells move from controls, but not whether they move together. We introduce perturbation stability, a geometric measure of the directional coherence of single-cell responses. Across five datasets and over 2,200 perturbations, stability tracks effect magnitude but departs from it in biologically revealing ways: pleiotropic regulators produce large but scattered cellular responses, while lineage-specific factors produce coherent shifts. After controlling for effect size, geometric instability is independently associated with elevated chaperone activation, and perturbations that are both coherent and highly stressed are systematically rare. The relationship persists in nonlinear foundation model embeddings, confirming it reflects biological geometry rather than projection artifact. Perturbation stability provides a complementary axis for screen evaluation and cell therapy quality control.

Abstract

Genome engineering has achieved remarkable sequence-level precision, yet predicting the transcriptomic state that a cell will occupy after perturbation remains an open problem. Single-cell CRISPR screens measure how far cells move from their unperturbed state, but this effect magnitude ignores a fundamental question: do the cells move together? Two perturbations with identical magnitude can produce qualitatively different outcomes if one drives cells coherently along a shared trajectory while the other scatters them across expression space. We introduce a geometric stability metric, Shesha, that quantifies the directional coherence of single-cell perturbation responses as the mean cosine similarity between individual cell shift vectors and the mean perturbation direction. Across five CRISPR datasets (2,200+ perturbations spanning CRISPRa, CRISPRi, and pooled screens), stability correlates strongly with effect magnitude (Spearman $\rho = 0.75$–0.97), with a calibrated cross-dataset correlation of 0.97. Crucially, discordant cases where the two metrics decouple expose regulatory architecture: pleiotropic master regulators such as CEBPA and GATA1 pay a "geometric tax," producing large but incoherent shifts, while lineage-specific factors such as KLF1 produce tightly coordinated responses. After controlling for magnitude, geometric instability is independently associated with elevated chaperone activation (HSPA5/BiP; $\rho_{\text{partial}} = -0.34$ and -0.21 across datasets), and the high-stability/high-stress quadrant is systematically depleted. The magnitude-stability relationship persists in scGPT foundation model embeddings, confirming it is a property of biological state space rather than linear projection. Perturbation stability provides a complementary axis for hit prioritization in screens, phenotypic quality control in cell manufacturing, and evaluation of in silico perturbation predictions.

The capacity to precisely edit genomes has outpaced our ability to predict the cellular consequences. CRISPR-Cas9 and its derivatives enable targeted modifications with unprecedented sequence-level accuracy (Doudna and Charpentier 2014; Jiang and Doudna 2017; Jinek et al. 2012), yet a cell can be edited exactly as intended and still drift toward an unintended fate. This gap between genetic precision and phenotypic predictability reflects three classes of failure that share a common feature. Off-target effects introduce unintended edits at genomically similar sites. On-target edits can trigger large deletions or chromothripsis invisible to standard sequencing (Kosicki, Tomberg, and Bradley 2018; Leibowitz et al. 2021). Most fundamentally, phenotypic heterogeneity is increasingly recognized as a defining challenge: two cells carrying the exact same edit often behave differently, one differentiating, one remaining stem-like (Replogle et al. 2022; Weinreb et al. 2020). This is not experimental noise. It reflects the initial position of each cell on the state manifold and the local geometry of that landscape.

These failures occur in cell state space, not sequence space. Current evaluation frameworks measure the syntax of the edit: indel rates, off-target cleavage, and sequence fidelity (Brinkman et al. 2014; Tsai et al. 2014). They answer the engineer's question: was the code changed correctly? They do not answer the biologist's question: is the resulting state stable? We have mastered the syntax of the genome (Doudna and Charpentier 2014; Jiang and Doudna 2017; Jinek et al. 2012). We remain largely blind to its semantics.

To resolve this gap, we must pivot from a sequence-centric view of perturbation biology to a geometric one. This is not a novel conceptual framework but a return to a foundational insight of developmental mechanics: the epigenetic landscape. When Conrad Waddington depicted cell development as a ball rolling down an undulating surface (Slack 2002; Waddington 1957), he was not merely offering an illustration. He was describing the topology of a dynamical system (Ferrell 2012). In this view, valleys are not metaphors; they are attractor basins, stable regions of state space where regulatory networks minimize the system's quasi-potential energy Enver et al. 2009; Fard et al. 2016; S. Huang 2009; Wang et al. 2011. Ridges separating valleys represent unstable intermediates where small perturbations can redirect trajectories. Gene regulatory networks are optimized not merely for specific expression patterns but for the stability of those patterns under perturbation (Kitano 2004; Siegal and Bergman 2002).

Modern single-cell genomics has transformed this topology from a theoretical construct into a measurable reality (Rand et al. 2021). We can now observe thousands of cells responding to the same genetic perturbation and ask not only how far they moved from their unperturbed state but how they moved relative to one another (Nadig et al. 2025; Norman et al. 2019; Replogle et al. 2022). Yet the standard analytical framework for single-cell CRISPR screens reduces this rich geometric information to a single summary: effect magnitude, the distance between the mean perturbed and control expression profiles. Recent work has begun to address within-perturbation heterogeneity, including harmonized benchmarking resources (Peidli et al. 2024) and single-cell perturbation response scores (Song et al. 2025), but no existing framework quantifies the directional coherence of a perturbation's population-level response. Two perturbations with identical magnitude can produce qualitatively different outcomes: one driving cells coherently along a shared transcriptomic trajectory, the other scattering them across expression space. Standard dimensionality reduction techniques (PCA, UMAP) compound this problem by projecting the high-dimensional state manifold onto flat coordinates, erasing the curvature that distinguishes deep attractors from shallow ridges (McInnes et al. 2018; Moon et al. 2019; Tsuyuzaki et al. 2020; Y. Zhou and Sharpee 2021). Two cell populations that appear phenotypically similar in a reduced projection may occupy positions on the manifold separated by high energetic barriers.

Here we introduce a geometric stability metric, Shesha, that quantifies the directional coherence of single-cell perturbation responses. For each perturbation, we compute the shift vector from the control centroid for every perturbed cell, then measure the mean cosine similarity between these individual shift vectors and the mean perturbation direction. This score, which we term perturbation stability (S_p), captures whether cells respond to a genetic intervention by moving together (high S_p, coherent) or scattering (low S_p, incoherent). The metric adapts the principle of geometric self-consistency from a general representational stability framework (Raju 2026a,b) to the specific context of perturbation biology.

We validate perturbation stability across five single-cell CRISPR datasets (Adamson et al. 2016; Dixit et al. 2016; Norman et al. 2019; Papalexi et al. 2021; Replogle et al. 2022) spanning activation (CRISPRa), interference (CRISPRi), and pooled screens, comprising over 2,200 perturbations. Stability correlates strongly with effect magnitude across all datasets (Spearman $\rho = 0.75–0.97$), but the cases where the two metrics decouple are the most informative: pleiotropic master regulators such as CEBPA and GATA1 produce large but geometrically incoherent shifts, while lineage-specific factors such as KLF1 produce tightly coordinated responses. This "geometric tax" on pleiotropy emerges without supervision and distinguishes regulatory architecture from the data alone. After controlling for effect size, geometric instability is independently associated with elevated chaperone activation (HSPA5/BiP), and the high-stability/high-stress quadrant is systematically depleted. The magnitude-stability relationship persists in nonlinear foundation model embeddings (scGPT; Cui et al. 2024), confirming it is a property of biological state space rather than an artifact of linear projection.

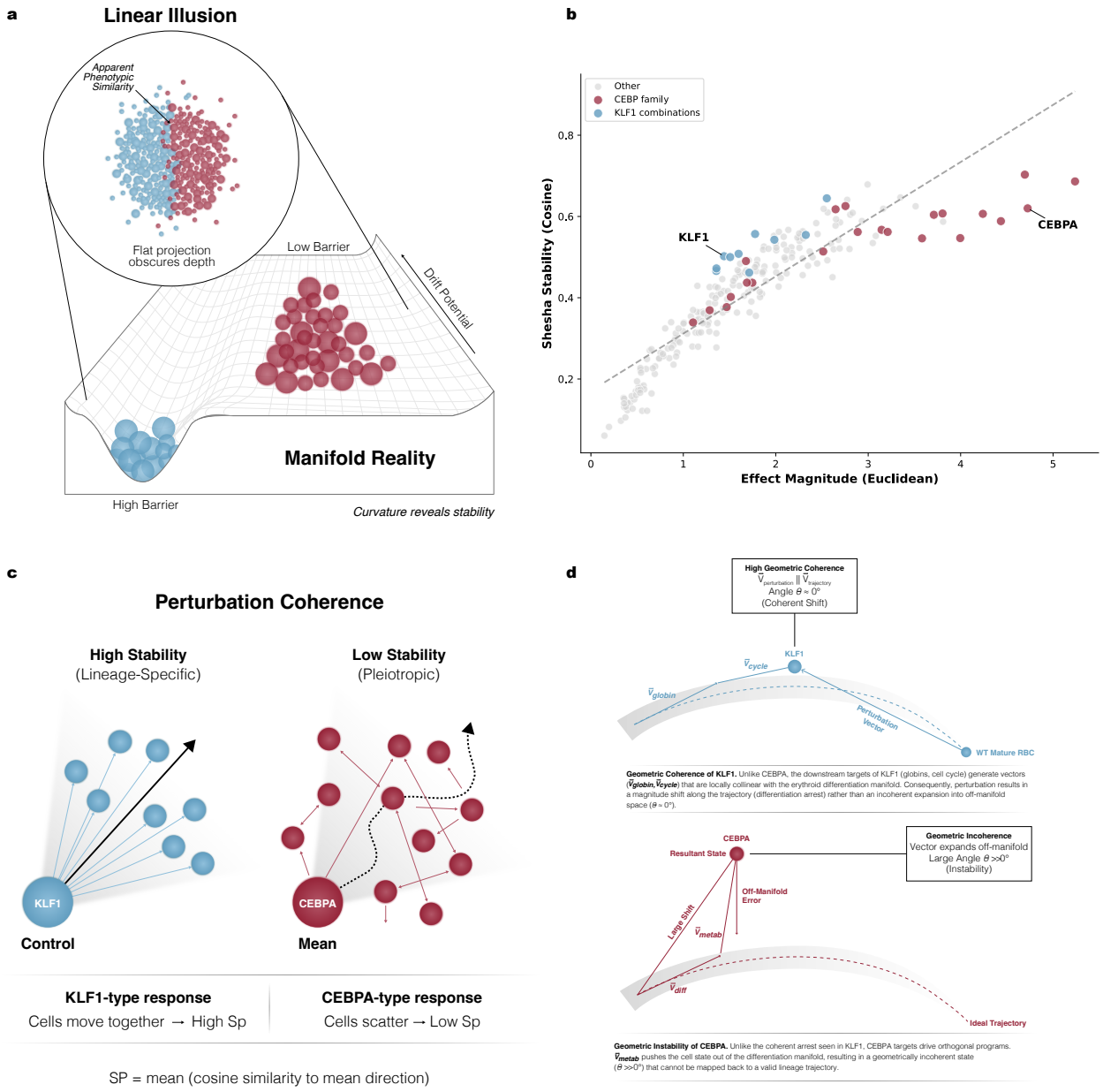

Figure 1: **The Geometric Tax: linear metrics obscure biological stability.** (**a**) Standard dimensionality reduction projects high-dimensional cell states onto a flat plane (Linear Illusion, inset), where two populations (blue, red) appear to overlap, suggesting similar phenotypes. Mapping these populations onto the underlying biological manifold (Manifold Reality) reveals distinct stability properties invisible to linear projections. The blue population occupies a deep valley (high barrier), representing a robust cell state resistant to perturbation. The red population sits on a shallow ridge (low barrier), representing an unstable state prone to drift. This stability difference constitutes the Geometric Tax of engineering cells into non-native configurations. (**b**) Stability versus magnitude for all 236 perturbations in Norman et al. (2019). CEBPA and perturbations involving CEBP family members (red) cluster below the regression line, indicating lower stability relative to their effect magnitude. KLF1 and its combinations (blue) cluster above. Dashed line: linear fit ($\rho = 0.953$). (**c**) Geometric stability quantified through perturbation coherence. High-stability perturbations (left, e.g., KLF1) produce shift vectors that align coherently, indicating cells move together along a shared trajectory toward the mean direction (solid arrow). Low-stability perturbations (right, e.g., CEBPA) scatter cells in divergent directions despite similar magnitude shifts, with the mean direction (dashed arc) representing dispersed cellular responses. The Shesha stability score (Sp) captures this distinction as the mean cosine similarity between individual shift vectors and the population mean. Together, panels a and b demonstrate how manifold curvature, invisible to linear projections, determines whether perturbations produce stable or fragile cellular states. (**d**) Geometric mechanism of perturbation coherence, illustrated schematically. Top: a lineage-specific perturbation (KLF1-type) produces shift vectors collinear with the differentiation manifold, resulting in coherent movement along an established trajectory (high S_p). Bottom: a pleiotropic perturbation (CEBPA-type) activates competing downstream programs whose shift vectors point in divergent directions, scattering cells off-manifold (low S_p).

Results

Quantifying geometric stability of perturbations

Single-cell CRISPR screens produce, for each perturbation, a population of cells whose transcriptomic profiles can be compared to unperturbed controls. The standard summary of this comparison is effect magnitude: the Euclidean distance between the mean perturbed and mean control expression profiles in a reduced-dimensional space. This scalar captures how far cells moved but discards all information about how they moved relative to one another.

We define perturbation stability S_p to recover this information. For a perturbation p applied to n_p cells, each perturbed cell i has a shift vector $\mathbf{d}_i = \mathbf{x}_i - \boldsymbol{\mu}_{\text{ctrl}}$, where $\mathbf{x}_i$ is the cell's position in PCA space and $\boldsymbol{\mu}_{\text{ctrl}}$ is the control centroid. The mean perturbation direction is $\bar{\mathbf{d}} = \frac{1}{n_p}\sum_i \mathbf{d}_i$, and perturbation stability is the mean cosine similarity between individual shift vectors and this mean direction:

$$S_p = \frac{1}{n_p}\sum_{i=1}^{n_p} \frac{\mathbf{d}_i \cdot \bar{\mathbf{d}}}{\|\mathbf{d}_i\|\,\|\bar{\mathbf{d}}\|} \tag{1}$$

A perturbation with S_p near 1 drives all cells in the same direction (high coherence); a perturbation with S_p near 0 scatters cells across expression space (low coherence). The metric adapts the principle of geometric self-consistency from the Shesha representational stability framework (Raju 2026a) to the specific context of perturbation biology.

We applied this metric to five publicly available single-cell CRISPR datasets spanning three perturbation modalities: CRISPRa activation (Norman et al. 2019, $n = 236$ perturbations in K562 cells), CRISPRi interference (Adamson et al. 2016, $n = 8$; Dixit et al. 2016, $n = 153$ in BMDCs; Replogle et al. 2022, $n = 1{,}832$ in K562), and a pooled screen (Papalexi et al. 2021, $n = 25$). All datasets were accessed via pertpy (Heumos et al. 2025) and processed through a standard pipeline (library-size normalization, log-transformation, 2,000 highly variable genes, PCA with 50 components; full details in **SI Appendix, Extended Methods**).

Stability tracks perturbation magnitude across modalities

Across all five datasets, perturbation stability correlates strongly and positively with effect magnitude (**Fig. 2**). The relationship is robust: Spearman correlations range from $\rho = 0.746$ in Dixit (95% CI $[0.641, 0.827]$) to $\rho = 0.985$ in Papalexi ($[0.939, 0.997]$), with the two largest datasets yielding $\rho = 0.953$ (Norman, $[0.934, 0.965]$) and $\rho = 0.970$ (Replogle, $[0.966, 0.972]$). Even Adamson, with only eight perturbations, shows $\rho = 0.929$ ($[0.407, 1.000]$). This consistency across CRISPRa and CRISPRi modalities, across cell types (K562, BMDCs, HeLa), and across screen scales (8 to 1,832 perturbations) indicates that the magnitude-stability relationship is a general property of perturbation geometry in single-cell expression space.

Table 1: Magnitude-stability correlation across five CRISPR datasets. Spearman ρ between effect magnitude (Euclidean distance) and perturbation stability (S_p, cosine coherence) in PCA space. Bootstrap 95% confidence intervals (10,000 resamples).

Dataset	Modality	n	Spearman ρ	95% CI
Norman 2019	CRISPRa	236	0.953	[0.934, 0.965]
Adamson 2016	CRISPRi	8	0.929	[0.407, 1.000]
Dixit 2016	CRISPRi	153	0.746	[0.641, 0.827]
Papalexi 2021	Pooled	25	0.985	[0.939, 0.997]
Replogle 2022	CRISPRi	1,832	0.970	[0.966, 0.972]
Pooled (z-scored)	—	2,254	0.968	[0.965, 0.971]

To confirm that this relationship generalizes beyond any single dataset, we z-scored magnitude and stability within each dataset and pooled all 2,254 perturbations. The calibrated cross-dataset correlation was $\rho = 0.968$ (95% CI $[0.965, 0.971]$; **Fig. 2f**). A linear mixed-effects model with dataset as a random effect confirmed that the dataset-level random-effect variance was near zero, and that magnitude was the dominant predictor of stability ($\beta = 0.168$, $[0.166, 0.170]$), accounting for approximately 11 times more variance than sample size ($\beta_{\text{n\_cells}} = -0.015$). The relationship was also robust to the choice of distance metric: Mahalanobis (whitened) and k-nearest neighbor matched controls produced consistent or stronger correlations (**SI Appendix, Robustness Analyses**).

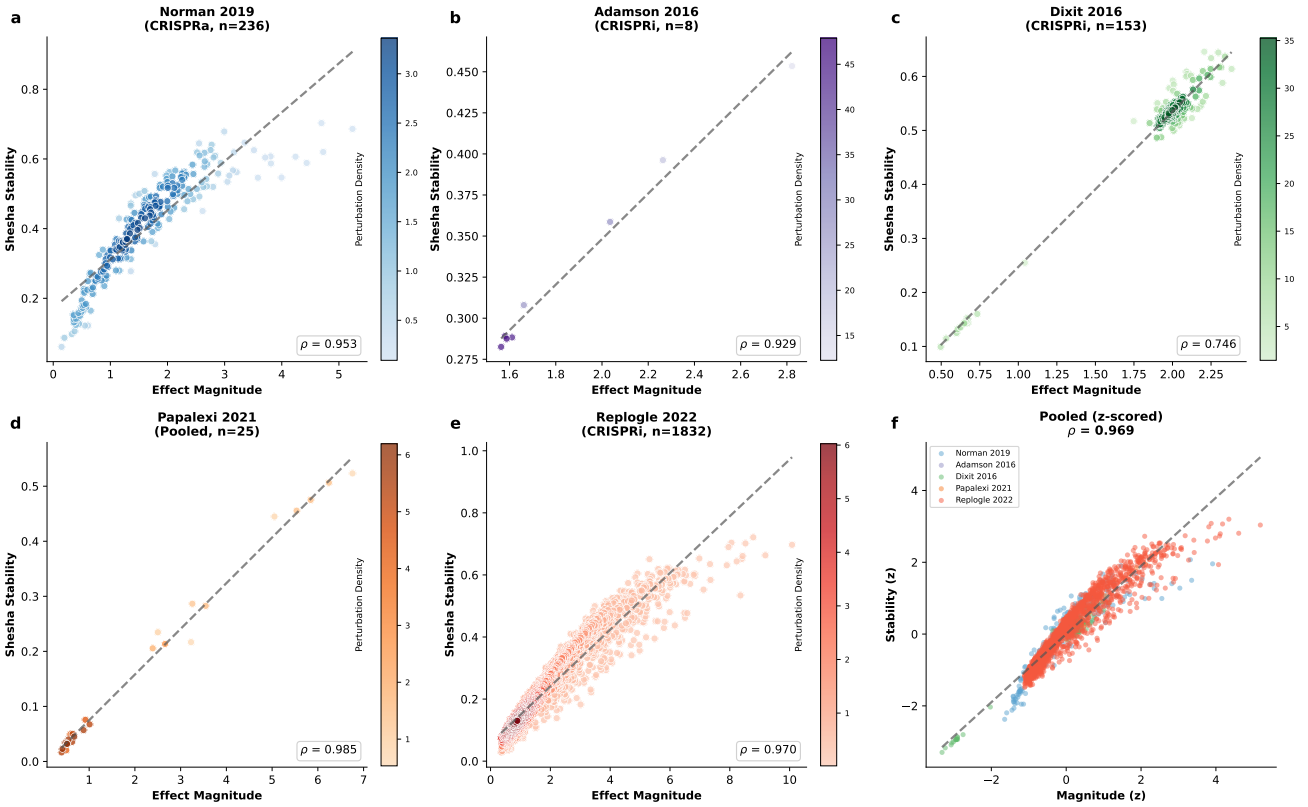

Figure 2: **Perturbation stability tracks effect magnitude across CRISPR modalities and cell types** (**a–e**) Effect magnitude (Euclidean distance, x-axis) vs perturbation stability (S_p, cosine coherence, y-axis) for each of five datasets: (**a**) Norman 2019 CRISPRa in K562 ($n = 236$, $\rho = 0.953$), (**b**) Adamson 2016 CRISPRi ($n = 8$, $\rho = 0.929$), (**c**) Dixit 2016 CRISPRi in BMDCs ($n = 153$, $\rho = 0.746$), (**d**) Papalexi 2021 pooled screen ($n = 25$, $\rho = 0.985$), (**e**) Replogle 2022 genome-scale CRISPRi in K562 ($n = 1{,}832$, $\rho = 0.970$). Dashed lines: linear regression. (**f**) Pooled cross-dataset scatter after z-score normalization within each dataset (calibrated $\rho = 0.968$, 95% CI $[0.965, 0.971]$). All Spearman correlations with bootstrap 95% CIs (10,000 resamples).

The strength of this correlation has a straightforward geometric interpretation: perturbations that push cells further from the control state tend to do so more coherently, because a large mean shift requires that individual shift vectors share a common direction. The interesting biology lies not in this expected correlation but in the cases where it breaks down.

Magnitude and stability decouple for pleiotropic regulators

Although magnitude and stability are strongly correlated, they are not redundant. To identify perturbations where the two metrics diverge, we computed discordance as the standardized residual from the magnitude-stability regression. Perturbations with large positive discordance produce large transcriptomic shifts but unexpectedly low geometric coherence; perturbations with large negative discordance are more coherent than their effect size would predict.

In the Norman CRISPRa dataset, the C/EBP transcription factor family shows consistently positive discordance (**Fig. 1d**). CEBPA, which activates downstream pathways spanning immune response, cell cycle control, metabolism, and lineage commitment(Avellino and Delwel 2017; Friedman 2007), drives perturbed cells far from controls but produces incoherent population-level responses. CEBPA and its combinatorial partners (CEBPA+JUN, CEBPA+CEBPB, CEBPA+CEBPE) cluster below the magnitude-stability regression line. KLF1, the erythroid-specific transcription factor whose targets are coordinated toward terminal red blood cell maturation (Miller and J J Bieker 1993; Pilon et al. 2008; Siatecka and James J. Bieker 2011; Tallack and Perkins 2010; Tallack, Whitington, et al. 2010), occupies the opposite end of the discordance spectrum. KLF1 and its combinations (KLF1+SET, BAK1+KLF1, FOXA1+KLF1) show moderate effect magnitudes but high stability, clustering above the regression line (**Fig. 1d**).

The same pattern emerges independently in the Replogle genome-scale CRISPRi screen (n=1,832 perturbations; **Fig. 6**) (Re-

plogle et al. 2022). Among the most discordant perturbations are GATA1 (master regulator of erythroid and megakaryocytic differentiation (Crispino and Weiss 2014); discordance =2.15, magnitude =10.08, stability =0.70), CHMP3 (ESCRT-III membrane remodeling complex (McCullough, Colf, and Sundquist 2013); discordance =2.14), and AQR (RNA helicase essential for spliceosome coupling (Hirose et al. 2006); discordance =1.80). At the concordant extreme, ribosome biogenesis factors (LSG1, ISG20L2, KRI1) produce tightly coherent shifts despite moderate effect sizes (Couté et al. 2008; Hedges, West, and Johnson 2005). Cell cycle regulators BUB3 (spindle assembly checkpoint (Taylor, Ha, and McKeon 1998)) and CENPW (centromere protein (Hori et al. 2008)) occupy the low-magnitude, low-stability corner of the scatter, indicating that even modest perturbations to essential mitotic genes produce geometrically incoherent cellular responses.

Two observations clarify the interpretation. Low geometric stability is not a proxy for cell cycle arrest: BUB3 (spindle assembly checkpoint) and CENPW (Hori et al. 2008; centromere protein) both show low stability in Replogle, but BLVRB (Wu et al. 2016; biliverdin reductase) shows high stability while cells continue cycling. Discordance quartile analysis reveals that the most discordant perturbations (Q4) have approximately three times the stability variance of the most concordant (Q1) in both Norman (SD = 0.202 vs 0.074) and Replogle (SD = 0.183 vs 0.097), while median cell counts are comparable across quartiles, ruling out a sample size confound. In every case, the shared feature of high-discordance perturbations is broad regulatory scope: transcription factors or essential complexes whose downstream targets span multiple functional programs.

Geometric instability is associated with cellular stress

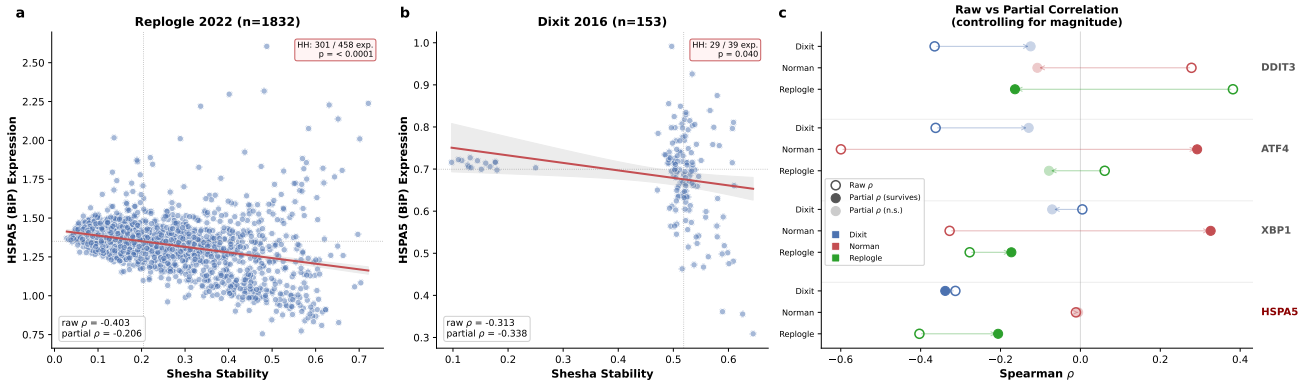

Figure 3: **Geometric instability is independently associated with chaperone activation** (**a**) Perturbation stability (x-axis) vs HSPA5 (BiP) expression (y-axis) in the Replogle 2022 CRISPRi dataset (n = 1,832). Red line: linear regression with 95% confidence interval (gray shading). Dotted lines: median splits defining quadrants. The high-stability/high-stress (HH) quadrant is depleted: 301 observed vs 458 expected under independence ($p < 10^{-18}$, one-sided binomial). Raw Spearman $\rho = -0.403$; partial $\rho = -0.206$ after controlling for magnitude. (**b**) Same analysis in Dixit 2016 CRISPRi (n = 153). HSPA5 shows the strongest partial correlation of any marker tested (partial $\rho = -0.338$, medium-large effect). HH quadrant depleted ($p = 0.040$). (**c**) Raw (open circles) vs partial (filled circles) Spearman correlations between stability and four stress markers (DDIT3, ATF4, XBP1, HSPA5) across three datasets, controlling for effect magnitude. Arrows connect raw to partial values. Filled circles at full opacity indicate correlations that survive magnitude control; faded circles indicate non-significant partial correlations. DDIT3 raw correlations collapse or flip sign after partialling (magnitude confound), while HSPA5 partial correlations persist in both CRISPRi datasets.

The geometric tax describes a structural property of perturbation responses, but does it have functional consequences? We hypothesized that geometrically incoherent perturbations, which scatter cells across expression space rather than guiding them along established trajectories, would be associated with elevated cellular stress. Cells pushed into off-manifold configurations that do not correspond to stable attractor states may activate homeostatic stress responses as they attempt to restore a viable gene expression program.

We tested this hypothesis by correlating perturbation stability with the mean expression of four stress and unfolded protein response (UPR) markers (DDIT3, ATF4, XBP1, HSPA5) across three datasets with sufficient sample size (Norman, Dixit, Replogle; Adamson excluded at $n = 8$). Because stability correlates with magnitude, and magnitude could independently affect stress gene expression, we report both raw Spearman correlations and partial correlations controlling for effect magnitude (**Fig. 3c**).

HSPA5 (BiP/GRP78), a canonical endoplasmic reticulum chaperone and marker of UPR activation (Oyadomari and Mori 2003), showed the most robust association with geometric instability. After controlling for magnitude, the partial correla-

tion between stability and HSPA5 expression remained significant in both CRISPRi datasets: Dixit (partial $\rho = -0.338$, 95% CI $[-0.506, -0.164]$, $p = 1.9\times10^{-5}$; medium-large effect) and Replogle (partial $\rho = -0.206$, $[-0.260, -0.152]$, $p = 5.2\times10^{-19}$; small-medium effect) (**Fig. 3a,b**). The direction is consistent: low geometric stability is associated with elevated chaperone activation, independent of effect size. The association was null in Norman CRISPRa (partial $\rho = -0.006$), which may reflect differences in how activation and interference perturbations engage stress pathways.

By contrast, DDIT3 (CHOP), a commonly used marker of the integrated stress response, was largely confounded by magnitude. Raw correlations were significant in all three datasets (Norman: $\rho = +0.278$; Dixit: $\rho = -0.365$; Replogle: $\rho = +0.382$), but after controlling for magnitude, DDIT3 failed to survive in Norman (partial $\rho = -0.108$, CI crosses zero) and Dixit (partial $\rho = -0.125$, CI crosses zero). The raw positive correlations in Norman and Replogle flipped negative after magnitude control, indicating that the raw signal was driven by larger perturbations having both higher stability and higher DDIT3 expression. DDIT3 survived only in Replogle (partial $\rho = -0.164$) with a small effect size (**Fig. 3c**). ATF4 and XBP1 showed mixed patterns across datasets, with sign inconsistencies that complicate interpretation (**SI Appendix, Robustness Analyses**).

We formalized the observation that geometric coherence and cellular stress are rarely coincident with a quadrant depletion test. Splitting perturbations at median stability and median stress expression, we tested whether the high-stability/high-stress (HH) quadrant was depleted relative to expectation under independence. For HSPA5, the HH quadrant was significantly depleted in both Dixit ($p = 0.040$, one-sided binomial) and Replogle ($p < 10^{-18}$), where only 301 perturbations occupied the HH quadrant compared to 458 expected under independence (odds ratio = 0.24) (**Fig. 3a,b**). This depletion was consistent for XBP1 but absent for DDIT3, further supporting the interpretation that DDIT3 is driven by magnitude rather than geometric instability per se. Full partial correlation tables, quadrant counts for all markers, and modality-stratified analyses are provided in **SI Appendix, Robustness Analyses**.

Taken together, these results indicate that geometric instability is not merely a structural descriptor but carries functional significance: perturbations that scatter cells incoherently across expression space are independently associated with activation of the chaperone stress response, suggesting that off-manifold cell states incur a measurable homeostatic cost.

Stability is a property of biological state space, not linear projection

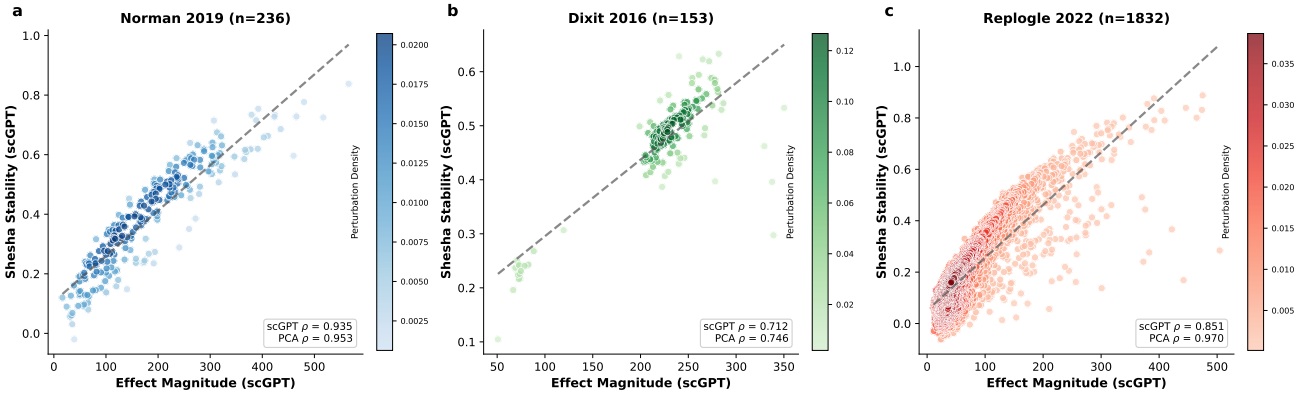

Figure 4: **The magnitude-stability relationship persists in nonlinear foundation model embeddings** Effect magnitude (x-axis) vs perturbation stability (y-axis) computed in scGPT "Whole Human" embeddings for three datasets: (**a**) Norman 2019 ($n = 236$, scGPT $\rho = 0.935$, PCA $\rho = 0.953$), (**b**) Dixit 2016 ($n = 153$, scGPT $\rho = 0.712$, PCA $\rho = 0.746$), (**c**) Replogle 2022 ($n = 1{,}832$, scGPT $\rho = 0.851$, PCA $\rho = 0.970$). Dashed lines: linear regression. The dataset rank order is preserved across embedding methods (Norman > Replogle > Dixit). scGPT correlations are consistently slightly lower than PCA, consistent with the nonlinear embedding resolving additional manifold structure that PCA collapses. The largest drop occurs in Replogle, the most diverse screen.

The results presented thus far rely on PCA embeddings, which project high-dimensional transcriptomic data onto a linear subspace. If the magnitude-stability relationship were an artifact of this linear projection, it would have limited biological significance. To test this, we replaced PCA with scGPT (Cui et al. 2024), a transformer-based foundation model pretrained on 33 million human cells that learns nonlinear representations of cell state. We computed perturbation stability and magnitude in scGPT embeddings for three datasets (Norman, Dixit, Replogle) using the "Whole Human" pretrained

checkpoint, with stability and magnitude computed identically to the PCA pipeline (full protocol in **SI Appendix, scGPT Validation Protocol**).

The magnitude-stability relationship persisted in every dataset (**Fig. 4**). Spearman correlations in scGPT embeddings were $\rho = 0.935$ for Norman (95% CI [0.911, 0.951]), $\rho = 0.712$ for Dixit ([0.585, 0.818]), and $\rho = 0.851$ for Replogle ([0.836, 0.865]), all highly significant ($p < 10^{-25}$). The dataset rank order was preserved across embedding methods (Norman > Replogle > Dixit), confirming that the framework captures real between-dataset variation rather than a ceiling effect.

The scGPT correlations were consistently slightly lower than their PCA counterparts (0.935 vs 0.953 for Norman; 0.712 vs 0.746 for Dixit; 0.851 vs 0.970 for Replogle). This is expected: a nonlinear embedding that resolves manifold structure flattened by PCA will introduce additional geometric complexity. Perturbations that appear coherent in a linear projection may reveal substructure, such as bifurcating trajectories or off-manifold curvature, in the learned embedding. The largest drop occurred in Replogle (0.970 to 0.851), consistent with the genome-scale screen containing the most diverse perturbation types where nonlinear geometry matters most.

These results establish that the magnitude-stability relationship is a property of biological state space rather than an artifact of the embedding method. They also suggest a practical criterion for evaluating foundation model representations: models that preserve the stability-magnitude structure have learned something geometrically faithful about cellular dynamics.

Combinatorial perturbations exhibit higher geometric stability

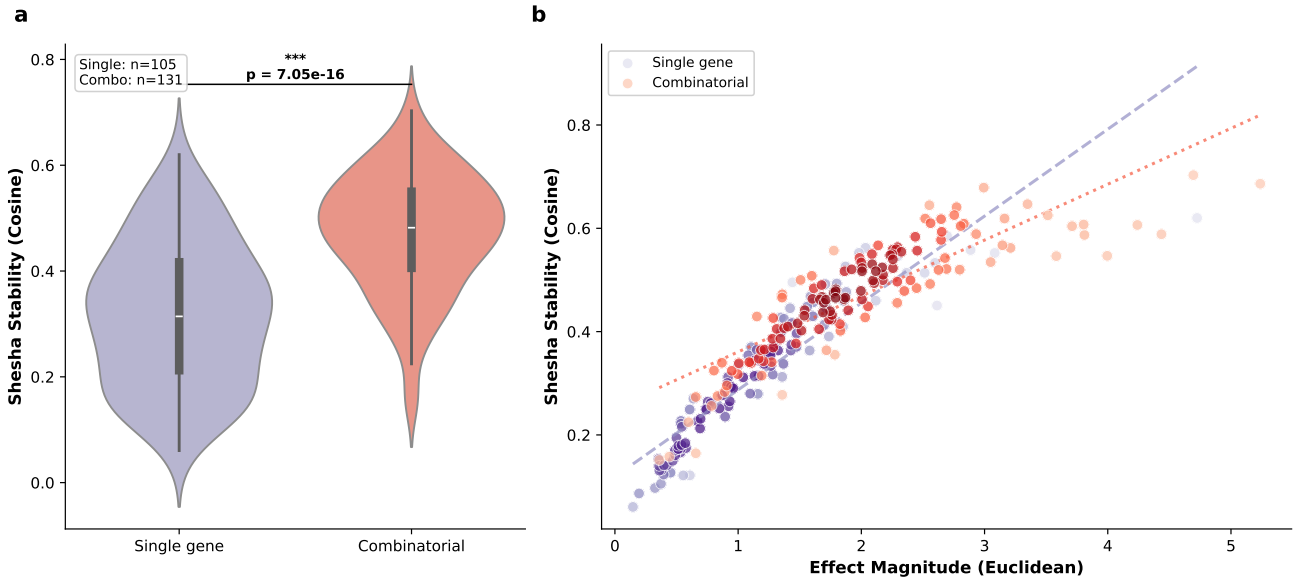

Figure 5: **Combinatorial perturbations exhibit higher geometric stability than single-gene perturbations** (**a**) Distribution of perturbation stability (S_p) for single-gene ($n = 105$) vs combinatorial ($n = 131$) perturbations in the Norman 2019 CRISPRa dataset. Combinatorial perturbations show significantly higher stability (Mann-Whitney U test). Boxes indicate interquartile range; internal line indicates median. (**b**) Magnitude-stability scatter colored by perturbation type. The magnitude-stability relationship holds within both categories with similar regression slopes, indicating that the higher stability of combinatorial perturbations is not explained by their larger effect magnitudes. Dashed line: single-gene regression; dotted line: combinatorial regression.

The Norman CRISPRa dataset contains both single-gene perturbations ($n = 105$) and combinatorial perturbations targeting two genes simultaneously ($n = 131$). Combinatorial perturbations showed significantly higher geometric stability than single-gene perturbations ($p < 10^{-9}$, Mann-Whitney U test; **Fig. 5a**), and this difference was not explained by combinatorial perturbations simply having larger effect magnitudes: the magnitude-stability relationship held within both categories, with similar regression slopes (**Fig. 5b**).

This observation is consistent with the Waddington landscape framework. When two genes are perturbed simultaneously, the combined intervention is more likely to engage an existing developmental trajectory, effectively pushing cells into a deeper attractor basin rather than scattering them off-manifold. Single-gene perturbations, by contrast, may activate only

one component of a multi-factor regulatory program, producing a partial shift that lacks the coordinated directionality of a full lineage transition. The higher coherence of combinatorial perturbations suggests that they more frequently align with canalized developmental programs, where the epigenetic landscape itself channels cells toward a stable endpoint.

Discussion

The geometric tax: a framework for interpreting perturbation coherence

The empirical results of this study converge on a single organizing principle. Perturbation magnitude and geometric stability are strongly correlated across datasets, modalities, and embedding methods, but they decouple in a biologically structured way: perturbations targeting pleiotropic regulators produce large but incoherent responses, while perturbations targeting lineage-specific factors produce coherent responses that align with established developmental trajectories. We propose the term "geometric tax" to describe this cost, measured in directional coherence, of activating broad regulatory programs.

The geometric tax is a consequence of regulatory network topology. When a transcription factor such as CEBPA (Friedman 2007) engages dozens of competing downstream pathways, no single cell can activate all target programs simultaneously. Each cell resolves the competition differently, producing a population-level incoherence that is invisible to standard effect-size metrics. By contrast, lineage-specific factors such as KLF1 (Siatecka and James J. Bieker 2011; Tallack, Whitington, et al. 2010), whose targets are functionally coordinated toward a single developmental outcome, maintain geometric coherence because the downstream programs are locally collinear on the state manifold. Two perturbations with identical magnitude can therefore occupy opposite ends of the stability spectrum, with dramatically different implications for reproducibility, phenotypic control, and therapeutic reliability (Li et al. 2016; J. X. Zhou, Brusch, and S. Huang 2011).

This framework connects to a deeper principle about biological robustness. Gene regulatory networks are optimized not merely for specific expression patterns but for the stability of those patterns under perturbation (Kitano 2004; Siegal and Bergman 2002). The geometric tax quantifies the cost of working against this optimization. In the language of Waddington's epigenetic landscape, perturbations into deep valleys, where canalized programs channel cells toward robust attractors, pay minimal tax (Davila-Velderrain, Martinez-Garcia, and Alvarez-Buylla 2015). Perturbations onto ridges or flat regions, where no strong attractor constrains the trajectory, pay heavily.

The association between geometric instability and chaperone activation provides functional evidence for this interpretation. Cells scattered into off-manifold configurations that do not correspond to stable attractor states activate the unfolded protein response (Lee 2005; Ron and Walter 2007), suggesting that geometric incoherence incurs a measurable homeostatic cost (Hetz 2012; Walter and Ron 2011). That this association survives after controlling for effect magnitude, and that the high-stability/high-stress quadrant is systematically depleted, indicates that the stress response is linked to the geometry of the perturbation rather than its strength alone. The geometric tax is therefore not merely structural but carries functional consequences: perturbations that pay the tax push cells into states that the cell itself recognizes as aberrant.

The tax is invisible to standard evaluation frameworks. Current metrics for perturbation screens, cell therapy manufacturing, and regulatory assessment focus on effect magnitude, marker expression, and sequence fidelity (Bravery et al. 2013; Simon et al. 2024). These capture the syntax of the intervention but miss its geometric semantics. A perturbation that scores well on all conventional metrics may nonetheless produce a geometrically incoherent cell population poised to drift toward unintended fates (Fraietta et al. 2018; Morris et al. 2014). Perturbation stability provides the missing axis for detecting this failure mode before it manifests clinically.

The alignment problem in cell engineering

The geometric tax framework addresses what might be called an alignment problem in biology: ensuring that engineered cells not only reach an intended expression profile but occupy a state that is dynamically stable. Current evaluation frameworks for cell therapies and engineered cell products focus on sequence fidelity (was the edit correct?) and marker expression (does the cell express the right genes?) (Vormittag et al. 2018). These metrics correspond to the syntax of the intervention. They do not assess whether the resulting cell state is geometrically coherent, whether it sits in a deep attractor basin, or whether it is poised to drift toward an unintended fate (Fraietta et al. 2018; Morris et al. 2014).

The clinical consequences of this gap are increasingly recognized. CAR-T cells optimized for tumor killing may find a local minimum in T cell exhaustion (Fraietta et al. 2018; Philip et al. 2017; Weber et al. 2021). Similarly, iPSC-derived beta cells may appear fully differentiated by marker expression, yet occupy a shallow basin that permits rapid de-differentiation upon transplantation (Lipsitz, Timmins, and Zandstra 2016; Nair et al. 2019; Talchai et al. 2012). In both cases, the failure lies not in the genetic intervention itself, but in the stability of the resulting cell state. Geometric stability provides a complementary axis for evaluating these outcomes: perturbations with a high S_p indicate that the intended phenotype coincides with a deep attractor, ensuring the cell product is fundamentally less likely to drift.

Practical applications

Three domains stand to benefit immediately from incorporating geometric stability into existing workflows. First, in high-throughput screening, stability can serve as a critical secondary ranking criterion alongside traditional effect size. This allows researchers to prioritize perturbations that produce replicable, coherent phenotypic shifts over those that generate superficially large, yet highly heterogeneous responses. Practically, perturbations that pay a high geometric tax (low S_p) are inherently poised to produce variable outcomes across experimental replicates, even when their mean effect sizes appear promising.

In cell manufacturing, stability offers a measure of phenotypic robustness that complements marker-based quality control (Bravery et al. 2013; Galipeau et al. 2016). A cell product with high stability occupies a coherent region of expression space, suggesting robust attractor occupancy. A product with low stability, even if it passes marker-based criteria, may be balanced on a flat region of the manifold where small perturbations could redirect trajectories. Stability could flag such products before clinical deployment.

In regulatory evaluation, geometric stability addresses a failure mode that current genotoxicity and potency assays cannot access (Salmikangas et al. 2023; Simon et al. 2024). The FDA's evolving framework for cell therapy evaluation emphasizes the continuous refinement of Critical Quality Attributes (CQAs) (Lipsitz, Timmins, and Zandstra 2016). Incorporating geometric stability as a formal CQA would provide a vital dynamical complement to existing static molecular and functional assessments.

Foundation models and in silico perturbation prediction

The validation of the magnitude-stability relationship in scGPT embeddings has implications beyond methodological robustness. Foundation models for single-cell biology, including scGPT (Cui et al. 2024), Geneformer (Theodoris et al. 2023), and UCE (Rosen et al. 2023), learn implicit representations of cell state geometry. Our results suggest that preservation of the stability-magnitude structure provides a necessary, though not sufficient, criterion for evaluating whether these learned representations capture biologically meaningful geometry.

For in silico perturbation prediction tools such as GEARS (Roohani, K. Huang, and Leskovec 2023), CellOracle (Kamimoto et al. 2023), and CPA (Lotfollahi et al. 2023) geometric stability offers a complementary quality metric. A predicted cell state with high magnitude shift but low predicted stability should be treated with skepticism: it may represent a computationally plausible but biologically unstable configuration, a "hallucinatory intermediate" that no real cell would occupy for long. Conversely, predictions that maintain high coherence are more likely to correspond to viable attractor states.

Practical considerations for implementing perturbation stability

The perturbation stability framework is implemented in the open-source Python package shesha-geometry, available on PyPI and designed for integration with standard single-cell analysis workflows built on scanpy (Wolf, Angerer, and Theis 2018) and AnnData (Virshup, Bredikhin, et al. 2023; Virshup, Rybakov, et al. 2024). The package provides functions for computing perturbation stability (S_p), effect magnitude, discordance scores, and bootstrap confidence intervals from any AnnData object containing a perturbation label and a control condition.

Computing S_p requires two choices: the embedding space and the distance metric. For the embedding space, we recommend PCA with 50 components as a default, consistent with standard single-cell preprocessing pipelines. Our results demonstrate that the magnitude-stability relationship is robust to the choice of embedding (PCA vs scGPT), but users working with foundation model embeddings or alternative dimensionality reduction methods can supply any cell-by-feature matrix. For the distance metric, Euclidean distance in PCA space serves as the default for both magnitude and

the shift vectors underlying S_p. Mahalanobis (whitened) and k-nearest neighbor matched controls produced consistent or stronger correlations in our benchmarks (**SI Appendix, Robustness Analyses**), and are available as options in the package.

Control group identification is handled by a multi-stage matching protocol that accommodates the heterogeneous labeling conventions across perturbation datasets (exact match, delimiter-aware regex, substring matching). Dataset-specific handling, such as pooling non-targeting guides in Papalexi or cleaning Replogle label prefixes, is documented in the package and in **SI Appendix, Extended Methods**. We recommend a minimum of 50 cells per perturbation for stable estimates; perturbations with fewer cells will produce wide bootstrap confidence intervals, as illustrated by the Adamson dataset ($n = 8$ perturbations, CI $[0.407, 1.000]$).

The package also provides utilities for discordance analysis (standardized residuals from the magnitude-stability regression), stress marker correlation (raw and partial, controlling for magnitude), and quadrant depletion tests. Our approach is complementary to recent methods for quantifying single-cell perturbation heterogeneity. The perturbation-response score (PS) of Song et al. (Song et al. 2025) estimates the strength of the perturbation effect for each individual cell, identifying which cells within a perturbation responded strongly. Perturbation stability (S_p) addresses a different question: among the cells that did respond, did they move in the same direction? A perturbation could have uniformly high PS values (all cells strongly affected) yet low S_p (each cell affected differently), as we observe for pleiotropic regulators like CEBPA. The two metrics capture orthogonal aspects of perturbation response and could be used jointly to distinguish coherent responders from scattered ones. We anticipate that perturbation stability will be most useful as a complement to existing effect-size metrics rather than a replacement, providing an orthogonal axis for evaluating screen hits, comparing perturbation modalities, and assessing the geometric coherence of in silico predictions.

Limitations

Several limitations should guide the interpretation of these results. First, PCA serves as the primary embedding throughout, though the scGPT validation mitigates concerns regarding linear projection artifacts. Future applications of manifold-aware methods, such as diffusion maps or PHATE (Moon et al. 2019), may reveal additional non-linear structure. Second, the Adamson dataset (Adamson et al. 2016) (n=8) provides limited statistical power, as reflected in its wide bootstrap confidence intervals. Third, the stress-stability association remains correlative rather than causal: we have not experimentally demonstrated that geometric incoherence explicitly causes stress activation, only that the two co-occur robustly after controlling for magnitude. Fourth, S_p operates as a global metric that summarizes each perturbation as a single scalar; it does not capture subpopulation-level structure, bifurcating responses, or dose-dependent heterogeneity within a single perturbation. Fifth, our analysis operates at the gene level rather than the guide level, meaning that unmeasured guide-level variation in perturbation efficiency could contribute to apparent incoherence. Finally, the association between geometric stability and alternative stress markers beyond HSPA5 exhibited heterogeneity across datasets. This indicates that the stress-stability relationship is heavily marker-specific and pathway-dependent, highlighting a need for further targeted studies to map these specific stress-response topologies.

Despite these limitations, the consistency of the magnitude-stability relationship across five datasets, two perturbation modalities, multiple cell types, and two fundamentally different embedding methods suggests that geometric stability captures a robust, universal property of how cells respond to genetic perturbation. The geometric tax framework provides a complementary axis—orthogonal to standard effect magnitude—for evaluating perturbation screens, assessing engineered cell product quality, and characterizing the deeper regulatory architecture that shapes cellular responses to intervention.

Materials and Methods

Datasets

Five single-cell CRISPR perturbation datasets were analyzed: Norman et al. 2019 (CRISPRa, K562, $n = 236$ perturbations), Adamson et al. 2016 (CRISPRi, $n = 8$), Dixit et al. 2016 (CRISPRi, BMDCs, $n = 153$), Papalexi et al. 2021 (pooled screen, $n = 25$), and Replogle et al. 2022 (genome-scale CRISPRi, K562, $n = 1{,}832$). All datasets were accessed via pertpy (Heumos et al. 2025). Each dataset was preprocessed independently using a standard scanpy pipeline: library-size normalization, log1p transformation, selection of the top 2,000 highly variable genes, and PCA with 50 components. Full preprocessing

details, including control group identification and dataset-specific handling, are provided in **SI Appendix, Extended Methods**.

Perturbation stability and effect magnitude

For each perturbation p with n_p cells, the shift vector for cell i is $\mathbf{d}_i = \mathbf{x}_i - \boldsymbol{\mu}_{\text{ctrl}}$, where $\boldsymbol{\mu}_{\text{ctrl}}$ is the control centroid in PCA space. Perturbation stability is the mean cosine similarity between individual shift vectors and the mean perturbation direction (Eq. 1). Effect magnitude is the Euclidean norm of the mean shift vector. Discordance is the standardized residual from the magnitude-stability linear regression. The metric adapts the Shesha geometric stability framework (Raju 2026a) to perturbation biology. Robustness to distance metric choice (Euclidean, Mahalanobis, k-NN) is confirmed in **SI Appendix, Robustness Analyses**.

Statistical analysis

All confidence intervals were computed via bootstrap resampling (10,000 iterations, seed 320, percentile method). Cross-dataset generalization was assessed with a linear mixed-effects model (dataset as random effect; fixed effects: magnitude, spread, sample size). Partial correlations between stability and stress markers were computed controlling for effect magnitude. Quadrant depletion was tested with a one-sided binomial test against the null expectation under independence of median-split categories. All p-values are two-sided unless otherwise noted. Full model specifications are provided in **SI Appendix, Mixed-Effects Model**.

scGPT validation

Cell embeddings were generated using the scGPT "Whole Human" pretrained checkpoint (Cui et al. 2024) on raw counts (not log-normalized) from Norman, Dixit, and Replogle datasets. Stability and magnitude were computed identically to the PCA pipeline. Full embedding protocol, including deterministic mode settings and batch parameters, is provided in **SI Appendix, scGPT Validation Protocol**.

Code

All analyses reported in this paper, including figure generation code, are available at `https://github.com/prashantcraju/geometric-stability-crispr`.

Acknowledgments

We thank Padma K. and Annapoorna Raju for generously supporting the computational resources used in this work. We thank the many institutions and individuals whose open-source datasets, frameworks, and models were used in our work. The authors acknowledge the use of large language models (specifically the GPT, Claude, and Gemini families) to assist with code debugging and text polishing. All hypotheses, experimental designs, analyses, and interpretations were independently formulated and verified by the authors, and the authors assume full responsibility for all content and claims in this work.

References

[1] Britt Adamson, Thomas M. Norman, Marco Jost, Min Y. Cho, James K. Nuñez, Yuwen Chen, Jacqueline E. Villalta, Luke A. Gilbert, Max A. Horlbeck, Marco Y. Hein, Ryan A. Pak, Andrew N. Gray, Carol A. Gross, Atray Dixit, Oren Parnas, Aviv Regev, and Jonathan S. Weissman. "A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response". In: *Cell* 167.7 (2016), 1867–1882.e21. ISSN: 0092-8674. DOI: 10.1016/j.cell.2016.11.048.

[2] Roberto Avellino and Ruud Delwel. "Expression and regulation of C/EBPα in normal myelopoiesis and in malignant transformation". In: *Blood* 129.15 (2017), pp. 2083–2091. ISSN: 1528-0020. DOI: 10.1182/blood-2016-09-687822.

[3] Christopher A. Bravery, Jessica Carmen, Timothy Fong, Wanda Oprea, Karin H. Hoogendoorn, Juliana Woda, Scott R. Burger, Jon A. Rowley, Mark L. Bonyhadi, and Wouter Van't Hof. "Potency assay development for cellular therapy products: an ISCT review of the requirements and experiences in the industry". In: *Cytotherapy* 15.1 (Jan. 2013), 9–19.e9. ISSN: 1465-3249. DOI: 10.1016/j.jcyt.2012.10.008. URL: http://dx.doi.org/10.1016/j.jcyt.2012.10.008.

[4] Eva K. Brinkman, Tao Chen, Mario Amendola, and Bas van Steensel. "Easy quantitative assessment of genome editing by sequence trace decomposition". In: *Nucleic Acids Research* 42.22 (Oct. 2014), e168–e168. ISSN: 0305-1048. DOI: 10.1093/nar/gku936. URL: http://dx.doi.org/10.1093/nar/gku936.

[5] Yohann Couté, Karine Kindbeiter, Stéphane Belin, Régis Dieckmann, Laurent Duret, Laurent Bezin, Jean-Charles Sanchez, and Jean-Jacques Diaz. "ISG20L2, a Novel Vertebrate Nucleolar Exoribonuclease Involved in Ribosome Biogenesis". In: *Molecular & Cellular Proteomics* 7.3 (Mar. 2008), pp. 546–559. ISSN: 1535-9476. DOI: 10.1074/mcp.m700510-mcp200. URL: http://dx.doi.org/10.1074/mcp.M700510-MCP200.

[6] John D. Crispino and Mitchell J. Weiss. "Erythro-megakaryocytic transcription factors associated with hereditary anemia". In: *Blood* 123.20 (May 2014), pp. 3080–3088. ISSN: 1528-0020. DOI: 10.1182/blood-2014-01-453167. URL: http://dx.doi.org/10.1182/blood-2014-01-453167.

[7] Haotian Cui, Chloe Wang, Hassaan Maan, Kuan Pang, Fengning Luo, Nan Duan, and Bo Wang. "scGPT: toward building a foundation model for single-cell multi-omics using generative AI". In: *Nature Methods* 21.8 (Feb. 2024), pp. 1470–1480. ISSN: 1548-7105. DOI: 10.1038/s41592-024-02201-0.

[8] Jose Davila-Velderrain, Juan C. Martinez-Garcia, and Elena R. Alvarez-Buylla. "Modeling the epigenetic attractors landscape: toward a post-genomic mechanistic understanding of development". In: *Frontiers in Genetics* 6 (Apr. 2015). ISSN: 1664-8021. DOI: 10.3389/fgene.2015.00160. URL: http://dx.doi.org/10.3389/fgene.2015.00160.

[9] Atray Dixit, Oren Parnas, Biyu Li, Jenny Chen, Charles P. Fulco, Livnat Jerby-Arnon, Nemanja D. Marjanovic, Danielle Dionne, Tyler Burks, Raktima Raychowdhury, Britt Adamson, Thomas M. Norman, Eric S. Lander, Jonathan S. Weissman, Nir Friedman, and Aviv Regev. "Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens". In: *Cell* 167.7 (2016), 1853–1866.e17. ISSN: 0092-8674. DOI: 10.1016/j.cell.2016.11.038.

[10] Jennifer A. Doudna and Emmanuelle Charpentier. "The new frontier of genome engineering with CRISPR-Cas9". In: *Science* 346.6213 (Nov. 2014). ISSN: 1095-9203. DOI: 10.1126/science.1258096. URL: http://dx.doi.org/10.1126/science.1258096.

[11] Tariq Enver, Martin Pera, Carsten Peterson, and Peter W. Andrews. "Stem Cell States, Fates, and the Rules of Attraction". In: *Cell Stem Cell* 4.5 (May 2009), pp. 387–397. ISSN: 1934-5909. DOI: 10.1016/j.stem.2009.04.011. URL: http://dx.doi.org/10.1016/j.stem.2009.04.011.

[12] Atefeh Taherian Fard, Sriganesh Srihari, Jessica C Mar, and Mark A Ragan. "Not just a colourful metaphor: modelling the landscape of cellular development using Hopfield networks". In: *npj Systems Biology and Applications* 2.1 (Feb. 2016). ISSN: 2056-7189. DOI: 10.1038/npjsba.2016.1. URL: http://dx.doi.org/10.1038/npjsba.2016.1.

[13] James E. Ferrell. "Bistability, Bifurcations, and Waddington's Epigenetic Landscape". In: *Current Biology* 22.11 (June 2012), R458–R466. ISSN: 0960-9822. DOI: 10.1016/j.cub.2012.03.045. URL: http://dx.doi.org/10.1016/j.cub.2012.03.045.

[14] Joseph A. Fraietta, Simon F. Lacey, Elena J. Orlando, Iulian Pruteanu-Malinici, Mercy Gohil, Stefan Lundh, Alina C. Boesteanu, Yan Wang, Roddy S. O'Connor, Wei-Ting Hwang, Edward Pequignot, David E. Ambrose, Changfeng Zhang, Nicholas Wilcox, Felipe Bedoya, Corin Dorfmeier, Fang Chen, Lifeng Tian, Harit Parakandi, Minnal Gupta, Regina M. Young, F. Brad Johnson, Irina Kulikovskaya, Li Liu, Jun Xu, Sadik H. Kassim, Megan M. Davis, Bruce L. Levine, Noelle V. Frey, Donald L. Siegel, Alexander C. Huang, E. John Wherry, Hans Bitter, Jennifer L. Brogdon, David L. Porter, Carl H. June, and J. Joseph Melenhorst. "Determinants of response and resistance to CD19 chimeric antigen receptor (CAR) T cell therapy of chronic lymphocytic leukemia". In: *Nature Medicine* 24.5 (Apr. 2018), pp. 563–571. ISSN: 1546-170X. DOI: 10.1038/s41591-018-0010-1. URL: http://dx.doi.org/10.1038/s41591-018-0010-1.

[15] Alan D. Friedman. "C/ebpα induces pu.1 and interacts with ap-1 and nf-κb to regulate myeloid development". In: *Blood Cells, Molecules, and Diseases* 39.3 (2007), pp. 340–343. ISSN: 1079-9796. DOI: 10.1016/j.bcmd.2007.06.010.

[16] Jacques Galipeau, Mauro Krampera, John Barrett, Francesco Dazzi, Robert J. Deans, Joost DeBruijn, Massimo Dominici, Willem E. Fibbe, Adrian P. Gee, Jeffery M. Gimble, Peiman Hematti, Mickey B.C. Koh, Katarina LeBlanc, Ivan Martin, Ian K. McNiece, Michael Mendicino, Steve Oh, Luis Ortiz, Donald G. Phinney, Valerie Planat, Yufang Shi, David F. Stroncek, Sowmya Viswanathan, Daniel J. Weiss, and Luc Sensebe. "International Society for Cellular Therapy perspective on immune functional assays for mesenchymal stromal cells as potency release criterion for

advanced phase clinical trials". In: *Cytotherapy* 18.2 (Feb. 2016), pp. 151–159. ISSN: 1465-3249. DOI: 10.1016/j.jcyt.2015.11.008. URL: http://dx.doi.org/10.1016/j.jcyt.2015.11.008.

[17] John Hedges, Matthew West, and Arlen W Johnson. "Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p". In: *The EMBO Journal* 24.3 (Jan. 2005), pp. 567–579. ISSN: 1460-2075. DOI: 10.1038/sj.emboj.7600547. URL: http://dx.doi.org/10.1038/sj.emboj.7600547.

[18] Claudio Hetz. "The unfolded protein response: controlling cell fate decisions under ER stress and beyond". In: *Nature Reviews Molecular Cell Biology* 13.2 (Jan. 2012), pp. 89–102. ISSN: 1471-0080. DOI: 10.1038/nrm3270. URL: http://dx.doi.org/10.1038/nrm3270.

[19] Lukas Heumos, Yuge Ji, Lilly May, Tessa D. Green, Stefan Peidli, Xinyue Zhang, Xichen Wu, Johannes Ostner, Antonia Schumacher, Karin Hrovatin, Michaela Müller, Faye Chong, Gregor Sturm, Alejandro Tejada, Emma Dann, Mingze Dong, Gonçalo Pinto, Mojtaba Bahrami, Ilan Gold, Sergei Rybakov, Altana Namsaraeva, Amir Ali Moinfar, Zihe Zheng, Eljas Roellin, Isra Mekki, Chris Sander, Mohammad Lotfollahi, Herbert B. Schiller, and Fabian J. Theis. "Pertpy: an end-to-end framework for perturbation analysis". In: *Nature Methods* 23.2 (Dec. 2025), pp. 350–359. ISSN: 1548-7105. DOI: 10.1038/s41592-025-02909-7. URL: http://dx.doi.org/10.1038/s41592-025-02909-7.

[20] Tetsuro Hirose, Takashi Ideue, Misato Nagai, Masatoshi Hagiwara, Mei-Di Shu, and Joan A. Steitz. "A Spliceosomal Intron Binding Protein, IBP160, Links Position-Dependent Assembly of Intron-Encoded Box C/D snoRNP to Pre-mRNA Splicing". In: *Molecular Cell* 23.5 (Sept. 2006), pp. 673–684. ISSN: 1097-2765. DOI: 10.1016/j.molcel.2006.07.011. URL: http://dx.doi.org/10.1016/j.molcel.2006.07.011.

[21] Tetsuya Hori, Miho Amano, Aussie Suzuki, Chelsea B. Backer, Julie P. Welburn, Yimin Dong, Bruce F. McEwen, Wei-Hao Shang, Emiko Suzuki, Katsuya Okawa, Iain M. Cheeseman, and Tatsuo Fukagawa. "CCAN Makes Multiple Contacts with Centromeric DNA to Provide Distinct Pathways to the Outer Kinetochore". In: *Cell* 135.6 (Dec. 2008), pp. 1039–1052. ISSN: 0092-8674. DOI: 10.1016/j.cell.2008.10.019. URL: http://dx.doi.org/10.1016/j.cell.2008.10.019.

[22] Sui Huang. "Reprogramming cell fates: reconciling rarity with robustness". In: *BioEssays* 31.5 (Apr. 2009), pp. 546–560. ISSN: 1521-1878. DOI: 10.1002/bies.200800189. URL: http://dx.doi.org/10.1002/bies.200800189.

[23] Fuguo Jiang and Jennifer A. Doudna. "CRISPR-Cas9 Structures and Mechanisms". In: *Annual Review of Biophysics* 46.1 (May 2017), pp. 505–529. ISSN: 1936-1238. DOI: 10.1146/annurev-biophys-062215-010822. URL: http://dx.doi.org/10.1146/annurev-biophys-062215-010822.

[24] Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. "A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity". In: *Science* 337.6096 (Aug. 2012), pp. 816–821. ISSN: 1095-9203. DOI: 10.1126/science.1225829. URL: http://dx.doi.org/10.1126/science.1225829.

[25] Kenji Kamimoto, Blerta Stringa, Christy M. Hoffmann, Kunal Jindal, Lilianna Solnica-Krezel, and Samantha A. Morris. "Dissecting cell identity via network inference and in silico gene perturbation". In: *Nature* 614.7949 (Feb. 2023), pp. 742–751. ISSN: 1476-4687. DOI: 10.1038/s41586-022-05688-9. URL: http://dx.doi.org/10.1038/s41586-022-05688-9.

[26] Hiroaki Kitano. "Biological robustness". In: *Nature Reviews Genetics* 5.11 (2004), pp. 826–837. ISSN: 1471-0064. DOI: 10.1038/nrg1471.

[27] Michael Kosicki, Kärt Tomberg, and Allan Bradley. "Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements". In: *Nature Biotechnology* 36.8 (July 2018), pp. 765–771. ISSN: 1546-1696. DOI: 10.1038/nbt.4192. URL: http://dx.doi.org/10.1038/nbt.4192.

[28] Amy S. Lee. "The ER chaperone and signaling regulator GRP78/BiP as a monitor of endoplasmic reticulum stress". In: *Methods* 35.4 (Apr. 2005), pp. 373–381. ISSN: 1046-2023. DOI: 10.1016/j.ymeth.2004.10.010. URL: http://dx.doi.org/10.1016/j.ymeth.2004.10.010.

[29] Mitchell L. Leibowitz, Stamatis Papathanasiou, Phillip A. Doerfler, Logan J. Blaine, Lili Sun, Yu Yao, Cheng-Zhong Zhang, Mitchell J. Weiss, and David Pellman. "Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing". In: *Nature Genetics* 53.6 (Apr. 2021), pp. 895–905. ISSN: 1546-1718. DOI: 10.1038/s41588-021-00838-7. URL: http://dx.doi.org/10.1038/s41588-021-00838-7.

[30] Qin Li, Anders Wennborg, Erik Aurell, Erez Dekel, Jie-Zhi Zou, Yuting Xu, Sui Huang, and Ingemar Ernberg. "Dynamics inside the cancer cell attractor reveal cell heterogeneity, limits of stability, and escape". In: *Proceedings of the National Academy of Sciences* 113.10 (Feb. 2016), pp. 2672–2677. ISSN: 1091-6490. DOI: 10.1073/pnas.1519210113. URL: http://dx.doi.org/10.1073/pnas.1519210113.

[31] Yonatan Y Lipsitz, Nicholas E Timmins, and Peter W Zandstra. "Quality cell therapy manufacturing by design". In: *Nature Biotechnology* 34.4 (Apr. 2016), pp. 393–400. ISSN: 1546-1696. DOI: 10.1038/nbt.3525. URL: http://dx.doi.org/10.1038/nbt.3525.

[32] Mohammad Lotfollahi, Anna Klimovskaia Susmelj, Carlo De Donno, Leon Hetzel, Yuge Ji, Ignacio L Ibarra, Sanjay R Srivatsan, Mohsen Naghipourfar, Riza M Daza, Beth Martin, Jay Shendure, Jose L McFaline-Figueroa, Pierre Boyeau, F Alexander Wolf, Nafissa Yakubova, Stephan Günnemann, Cole Trapnell, David Lopez-Paz, and Fabian J Theis. "Predicting cellular responses to complex perturbations in high-throughput screens". In: *Molecular Systems Biology* 19.6 (May 2023). ISSN: 1744-4292. DOI: 10.15252/msb.202211517. URL: http://dx.doi.org/10.15252/msb.202211517.

[33] John McCullough, Leremy A. Colf, and Wesley I. Sundquist. "Membrane Fission Reactions of the Mammalian ESCRT Pathway". In: *Annual Review of Biochemistry* 82.1 (June 2013), pp. 663–692. ISSN: 1545-4509. DOI: 10.1146/annurev-biochem-072909-101058. URL: http://dx.doi.org/10.1146/annurev-biochem-072909-101058.

[34] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. "UMAP: Uniform Manifold Approximation and Projection". In: *Journal of Open Source Software* 3.29 (Sept. 2018), p. 861. ISSN: 2475-9066. DOI: 10.21105/joss.00861. URL: http://dx.doi.org/10.21105/joss.00861.

[35] I J Miller and J J Bieker. "A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins." In: *Molecular and Cellular Biology* 13.5 (1993), pp. 2776–2786. ISSN: 1098-5549. DOI: 10.1128/mcb.13.5.2776.

[36] Kevin R. Moon, David van Dijk, Zheng Wang, Scott Gigante, Daniel B. Burkhardt, William S. Chen, Kristina Yim, Antonia van den Elzen, Matthew J. Hirn, Ronald R. Coifman, Natalia B. Ivanova, Guy Wolf, and Smita Krishnaswamy. "Visualizing structure and transitions in high-dimensional biological data". In: *Nature Biotechnology* 37.12 (Dec. 2019), pp. 1482–1492. ISSN: 1546-1696. DOI: 10.1038/s41587-019-0336-3. URL: http://dx.doi.org/10.1038/s41587-019-0336-3.

[37] Samantha A. Morris, Patrick Cahan, Hu Li, Anna M. Zhao, Adrianna K. San Roman, Ramesh A. Shivdasani, James J. Collins, and George Q. Daley. "Dissecting Engineered Cell Types and Enhancing Cell Fate Conversion via CellNet". In: *Cell* 158.4 (Aug. 2014), pp. 889–902. ISSN: 0092-8674. DOI: 10.1016/j.cell.2014.07.021. URL: http://dx.doi.org/10.1016/j.cell.2014.07.021.

[38] Ajay Nadig, Joseph M. Replogle, Angela N. Pogson, Mukundh Murthy, Steven A. McCarroll, Jonathan S. Weissman, Elise B. Robinson, and Luke J. O'Connor. "Transcriptome-wide analysis of differential expression in perturbation atlases". In: *Nature Genetics* 57.5 (Apr. 2025), pp. 1228–1237. ISSN: 1546-1718. DOI: 10.1038/s41588-025-02169-3. URL: http://dx.doi.org/10.1038/s41588-025-02169-3.

[39] Gopika G. Nair, Jennifer S. Liu, Holger A. Russ, Stella Tran, Michael S. Saxton, Richard Chen, Charity Juang, Meilan Li, Vinh Q. Nguyen, Simone Giacometti, Sapna Puri, Yuan Xing, Yong Wang, Gregory L. Szot, Jose Oberholzer, Anil Bhushan, and Matthias Hebrok. "Recapitulating endocrine cell clustering in culture promotes maturation of human stem-cell-derived β cells". In: *Nature Cell Biology* 21.2 (Feb. 2019), pp. 263–274. ISSN: 1476-4679. DOI: 10.1038/s41556-018-0271-4. URL: http://dx.doi.org/10.1038/s41556-018-0271-4.

[40] Thomas M Norman, Max A Horlbeck, Joseph M Replogle, Y Alexander Ge, Alex Xu, Marco Jost, Luke A Gilbert, and Jonathan S Weissman. "Exploring genetic interaction manifolds constructed from rich single-cell phenotypes". In: *Science* 365.6455 (2019), pp. 786–793. ISSN: 1095-9203. DOI: 10.1126/science.aax4438.

[41] S Oyadomari and M Mori. "Roles of CHOP/GADD153 in endoplasmic reticulum stress". In: *Cell Death & Differentiation* 11.4 (Dec. 2003), pp. 381–389. ISSN: 1476-5403. DOI: 10.1038/sj.cdd.4401373. URL: http://dx.doi.org/10.1038/sj.cdd.4401373.

[42] Efthymia Papalexi, Eleni P. Mimitou, Andrew W. Butler, Samantha Foster, Bernadette Bracken, William M. Mauck, Hans-Hermann Wessels, Yuhan Hao, Bertrand Z. Yeung, Peter Smibert, and Rahul Satija. "Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens". In: *Nature Genetics* 53.3 (2021), pp. 322–331. ISSN: 1546-1718. DOI: 10.1038/s41588-021-00778-2.

[43] Stefan Peidli, Tessa D. Green, Ciyue Shen, Torsten Gross, Joseph Min, Samuele Garda, Bo Yuan, Linus J. Schumacher, Jake P. Taylor-King, Debora S. Marks, Augustin Luna, Nils Blüthgen, and Chris Sander. "scPerturb: harmonized single-cell perturbation data". In: *Nature Methods* 21.3 (Jan. 2024), pp. 531–540. ISSN: 1548-7105. DOI: 10.1038/s41592-023-02144-y. URL: http://dx.doi.org/10.1038/s41592-023-02144-y.

[44] Mary Philip, Lauren Fairchild, Liping Sun, Ellen L. Horste, Steven Camara, Mojdeh Shakiba, Andrew C. Scott, Agnes Viale, Peter Lauer, Taha Merghoub, Matthew D. Hellmann, Jedd D. Wolchok, Christina S. Leslie, and Andrea Schietinger. "Chromatin states define tumour-specific T cell dysfunction and reprogramming". In: *Nature* 545.7655

(May 2017), pp. 452–456. ISSN: 1476-4687. DOI: 10.1038/nature22367. URL: http://dx.doi.org/10.1038/nature22367.

[45] Andre M. Pilon, Murat O. Arcasoy, Holly K. Dressman, Serena E. Vayda, Yelena D. Maksimova, Jose I. Sangerman, Patrick G. Gallagher, and David M. Bodine. "Failure of Terminal Erythroid Differentiation in EKLF-Deficient Mice Is Associated with Cell Cycle Perturbation and Reduced Expression of E2F2". In: *Molecular and Cellular Biology* 28.24 (2008), pp. 7394–7401. ISSN: 1098-5549. DOI: 10.1128/mcb.01087-08.

[46] Prashant C. Raju. "Geometric Stability: The Missing Axis of Representations". In: *arXiv preprint arXiv:2601.09173* (2026).

[47] Prashant C. Raju. *Shesha: Self-Consistency Metrics for Representational Stability*. 2026. DOI: 10.5281/zenodo.18227453. URL: https://doi.org/10.5281/zenodo.18227453.

[48] David A. Rand, Archishman Raju, Meritxell Sáez, Francis Corson, and Eric D. Siggia. "Geometry of gene regulatory dynamics". In: *Proceedings of the National Academy of Sciences* 118.38 (Sept. 2021). ISSN: 1091-6490. DOI: 10.1073/pnas.2109729118. URL: http://dx.doi.org/10.1073/pnas.2109729118.

[49] Joseph M. Replogle, Reuben A. Saunders, Angela N. Pogson, Jeffrey A. Hussmann, Alexander Lenail, Alina Guna, Lauren Mascibroda, Eric J. Wagner, Karen Adelman, Gila Lithwick-Yanai, Nika Iremadze, Florian Oberstrass, Doron Lipson, Jessica L. Bonnar, Marco Jost, Thomas M. Norman, and Jonathan S. Weissman. "Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq". In: *Cell* 185.14 (July 2022), 2559–2575.e28. ISSN: 0092-8674. DOI: 10.1016/j.cell.2022.05.013.

[50] David Ron and Peter Walter. "Signal integration in the endoplasmic reticulum unfolded protein response". In: *Nature Reviews Molecular Cell Biology* 8.7 (July 2007), pp. 519–529. ISSN: 1471-0080. DOI: 10.1038/nrm2199. URL: http://dx.doi.org/10.1038/nrm2199.

[51] Yusuf Roohani, Kexin Huang, and Jure Leskovec. "Predicting transcriptional outcomes of novel multigene perturbations with GEARS". In: *Nature Biotechnology* 42.6 (Aug. 2023), pp. 927–935. ISSN: 1546-1696. DOI: 10.1038/s41587-023-01905-6. URL: http://dx.doi.org/10.1038/s41587-023-01905-6.

[52] Yanay Rosen, Yusuf Roohani, Ayush Agrawal, Leon Samotorcan, Tabula Sapiens Consortium, Stephen R. Quake, and Jure Leskovec. "Universal Cell Embeddings: A Foundation Model for Cell Biology". In: *bioRxiv* (2023). DOI: 10.1101/2023.11.28.568918.

[53] Paula Salmikangas, Björn Carlsson, Christophe Klumb, Tatiana Reimer, and Steffen Thirstrup. "Potency testing of cell and gene therapy products". In: *Frontiers in Medicine* 10 (May 2023). ISSN: 2296-858X. DOI: 10.3389/fmed.2023.1190016. URL: http://dx.doi.org/10.3389/fmed.2023.1190016.

[54] Miroslawa Siatecka and James J. Bieker. "The multifunctional role of EKLF/KLF1 during erythropoiesis". In: *Blood* 118.8 (2011), pp. 2044–2054. ISSN: 1528-0020. DOI: 10.1182/blood-2011-03-331371.

[55] Mark L. Siegal and Aviv Bergman. "Waddington's canalization revisited: Developmental stability and evolution". In: *Proceedings of the National Academy of Sciences* 99.16 (June 2002), pp. 10528–10532. ISSN: 1091-6490. DOI: 10.1073/pnas.102303999. URL: http://dx.doi.org/10.1073/pnas.102303999.

[56] Carl G. Simon, Erich H. Bozenhardt, Christina M. Celluzzi, David Dobnik, Melanie L. Grant, Uma Lakshmipathy, Thiana Nebel, Linda Peltier, Anthony Ratcliffe, James L. Sherley, Glyn N. Stacey, Rouzbeh R. Taghizadeh, Eddie H. P. Tan, and Sandrine Vessillier. "Mechanism of action, potency and efficacy: considerations for cell therapies". In: *Journal of Translational Medicine* 22.1 (May 2024). ISSN: 1479-5876. DOI: 10.1186/s12967-024-05179-7. URL: http://dx.doi.org/10.1186/s12967-024-05179-7.

[57] Jonathan M. W. Slack. "Conrad Hal Waddington: the last Renaissance biologist?" In: *Nature Reviews Genetics* 3.11 (Nov. 2002), pp. 889–895. ISSN: 1471-0064. DOI: 10.1038/nrg933. URL: http://dx.doi.org/10.1038/nrg933.

[58] Bicna Song, Dingyu Liu, Weiwei Dai, Natalie F. McMyn, Qingyang Wang, Dapeng Yang, Adam Krejci, Anatoly Vasilyev, Nicole Untermoser, Anke Loregger, Dongyuan Song, Breanna Williams, Bess Rosen, Xiaolong Cheng, Lumen Chao, Hanuman T. Kale, Hao Zhang, Yarui Diao, Tilmann Bürckstümmer, Janet D. Siliciano, Jingyi Jessica Li, Robert F. Siliciano, Danwei Huangfu, and Wei Li. "Decoding heterogeneous single-cell perturbation responses". In: *Nature Cell Biology* 27.3 (Feb. 2025), pp. 493–504. ISSN: 1476-4679. DOI: 10.1038/s41556-025-01626-9. URL: http://dx.doi.org/10.1038/s41556-025-01626-9.

[59] Chutima Talchai, Shouhong Xuan, Hua V. Lin, Lori Sussel, and Domenico Accili. "Pancreatic β Cell Dedifferentiation as a Mechanism of Diabetic β Cell Failure". In: *Cell* 150.6 (Sept. 2012), pp. 1223–1234. ISSN: 0092-8674. DOI: 10.1016/j.cell.2012.07.029. URL: http://dx.doi.org/10.1016/j.cell.2012.07.029.

[60] Michael R. Tallack and Andrew C. Perkins. "KLF1 directly coordinates almost all aspects of terminal erythroid differentiation". In: *IUBMB Life* 62.12 (2010), pp. 886–890. ISSN: 1521-6551. DOI: 10.1002/iub.404.

[61] Michael R. Tallack, Tom Whitington, Wai Shan Yuen, Elanor N. Wainwright, Janelle R. Keys, Brooke B. Gardiner, Ehsan Nourbakhsh, Nicole Cloonan, Sean M. Grimmond, Timothy L. Bailey, and Andrew C. Perkins. "A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells". In: *Genome Research* 20.8 (2010), pp. 1052–1063. ISSN: 1088-9051. DOI: 10.1101/gr.106575.110.

[62] Stephen S. Taylor, Edward Ha, and Frank McKeon. "The Human Homologue of Bub3 Is Required for Kinetochore Localization of Bub1 and a Mad3/Bub1-related Protein Kinase". In: *The Journal of Cell Biology* 142.1 (July 1998), pp. 1–11. ISSN: 1540-8140. DOI: 10.1083/jcb.142.1.1. URL: http://dx.doi.org/10.1083/jcb.142.1.1.

[63] Christina V. Theodoris, Ling Xiao, Anant Chopra, Mark D. Chaffin, Zeina R. Al Sayed, Matthew C. Hill, Helene Mantineo, Elizabeth M. Brydon, Zexian Zeng, X. Shirley Liu, and Patrick T. Ellinor. "Transfer learning enables predictions in network biology". In: *Nature* 618.7965 (May 2023), pp. 616–624. ISSN: 1476-4687. DOI: 10.1038/s41586-023-06139-9.

[64] Shengdar Q Tsai, Zongli Zheng, Nhu T Nguyen, Matthew Liebers, Ved V Topkar, Vishal Thapar, Nicolas Wyvekens, Cyd Khayter, A John Iafrate, Long P Le, Martin J Aryee, and J Keith Joung. "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas9 nucleases". In: *Nature Biotechnology* 33.2 (Dec. 2014), pp. 187–197. ISSN: 1546-1696. DOI: 10.1038/nbt.3117. URL: http://dx.doi.org/10.1038/nbt.3117.

[65] Koki Tsuyuzaki, Hiroyuki Sato, Kenta Sato, and Itoshi Nikaido. "Benchmarking principal component analysis for large-scale single-cell RNA-sequencing". In: *Genome Biology* 21.1 (Jan. 2020). ISSN: 1474-760X. DOI: 10.1186/s13059-019-1900-3. URL: http://dx.doi.org/10.1186/s13059-019-1900-3.

[66] Isaac Virshup, Danila Bredikhin, Lukas Heumos, Giovanni Palla, Gregor Sturm, Adam Gayoso, Ilia Kats, Mikaela Koutrouli, Philipp Angerer, Volker Bergen, Pierre Boyeau, Maren Büttner, Gokcen Eraslan, David Fischer, Max Frank, Justin Hong, Michal Klein, Marius Lange, Romain Lopez, Mohammad Lotfollahi, Malte D. Luecken, Fidel Ramirez, Jeffrey Regier, Sergei Rybakov, Anna C. Schaar, Valeh Valiollah Pour Amiri, Philipp Weiler, Galen Xing, Bonnie Berger, Dana Pe'er, Aviv Regev, Sarah A. Teichmann, Francesca Finotello, F. Alexander Wolf, Nir Yosef, Oliver Stegle, and Fabian J. Theis. "The scverse project provides a computational ecosystem for single-cell omics data analysis". In: *Nature Biotechnology* 41.5 (Apr. 2023), pp. 604–606. ISSN: 1546-1696. DOI: 10.1038/s41587-023-01733-8. URL: http://dx.doi.org/10.1038/s41587-023-01733-8.

[67] Isaac Virshup, Sergei Rybakov, Fabian J. Theis, Philipp Angerer, and F. Alexander Wolf. "anndata: Access and store annotated data matrices". In: *Journal of Open Source Software* 9.101 (Sept. 2024), p. 4371. ISSN: 2475-9066. DOI: 10.21105/joss.04371. URL: http://dx.doi.org/10.21105/joss.04371.

[68] Philipp Vormittag, Rebecca Gunn, Sara Ghorashian, and Farlan S Veraitch. "A guide to manufacturing CAR T cell therapies". In: *Current Opinion in Biotechnology* 53 (Oct. 2018), pp. 164–181. ISSN: 0958-1669. DOI: 10.1016/j.copbio.2018.01.025. URL: http://dx.doi.org/10.1016/j.copbio.2018.01.025.

[69] Conrad Hal Waddington. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. London: George Allen & Unwin, 1957.

[70] Peter Walter and David Ron. "The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation". In: *Science* 334.6059 (Nov. 2011), pp. 1081–1086. ISSN: 1095-9203. DOI: 10.1126/science.1209038. URL: http://dx.doi.org/10.1126/science.1209038.

[71] Jin Wang, Kun Zhang, Li Xu, and Erkang Wang. "Quantifying the Waddington landscape and biological paths for development and differentiation". In: *Proceedings of the National Academy of Sciences* 108.20 (May 2011), pp. 8257–8262. ISSN: 1091-6490. DOI: 10.1073/pnas.1017017108. URL: http://dx.doi.org/10.1073/pnas.1017017108.

[72] Evan W. Weber, Kevin R. Parker, Elena Sotillo, Rachel C. Lynn, Hima Anbunathan, John Lattin, Zinaida Good, Julia A. Belk, Bence Daniel, Dorota Klysz, Meena Malipatlolla, Peng Xu, Malek Bashti, Sabine Heitzeneder, Louai Labanieh, Panayiotis Vandris, Robbie G. Majzner, Yanyan Qi, Katalin Sandor, Ling-Chun Chen, Snehit Prabhu, Andrew J. Gentles, Thomas J. Wandless, Ansuman T. Satpathy, Howard Y. Chang, and Crystal L. Mackall. "Transient rest restores functionality in exhausted CAR-T cells through epigenetic remodeling". In: *Science* 372.6537 (Apr. 2021). ISSN: 1095-9203. DOI: 10.1126/science.aba1786. URL: http://dx.doi.org/10.1126/science.aba1786.

[73] Caleb Weinreb, Alejo Rodriguez-Fraticelli, Fernando D. Camargo, and Allon M. Klein. "Lineage tracing on transcriptional landscapes links state to fate during differentiation". In: *Science* 367.6479 (Feb. 2020). ISSN: 1095-9203. DOI: 10.1126/science.aaw3381. URL: http://dx.doi.org/10.1126/science.aaw3381.

[74] F. Alexander Wolf, Philipp Angerer, and Fabian J. Theis. "SCANPY: large-scale single-cell gene expression data analysis". In: *Genome Biology* 19.1 (2018). ISSN: 1474-760X. DOI: 10.1186/s13059-017-1382-0.

[75] Song Wu, Zongdong Li, Dmitri V. Gnatenko, Beibei Zhang, Lu Zhao, Lisa E. Malone, Nedialka Markova, Timothy J. Mantle, Natasha M. Nesbitt, and Wadie F. Bahou. "BLVRB redox mutation defines heme degradation in a metabolic

pathway of enhanced thrombopoiesis in humans". In: *Blood* 128.5 (Aug. 2016), pp. 699–709. ISSN: 1528-0020. DOI: 10.1182/blood-2016-02-696997. URL: http://dx.doi.org/10.1182/blood-2016-02-696997.

[76] Joseph Xu Zhou, Lutz Brusch, and Sui Huang. "Predicting Pancreas Cell Fate Decisions and Reprogramming with a Hierarchical Multi-Attractor Model". In: *PLoS ONE* 6.3 (Mar. 2011). Ed. by Kumar Selvarajoo, e14752. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0014752. URL: http://dx.doi.org/10.1371/journal.pone.0014752.

[77] Yuansheng Zhou and Tatyana O. Sharpee. "Hyperbolic geometry of gene expression". In: *iScience* 24.3 (Mar. 2021), p. 102225. ISSN: 2589-0042. DOI: 10.1016/j.isci.2021.102225. URL: http://dx.doi.org/10.1016/j.isci.2021.102225.

Extended Methods

Datasets and preprocessing

Five single-cell CRISPR perturbation datasets were accessed via the pertpy Python package (version 1.0.4) (Heumos et al. 2025). Table 2 summarizes the datasets. Each dataset was preprocessed independently to prevent batch effects. The pipeline for each dataset consisted of:

1. Quality filtering: cells with fewer than 100 detected genes were removed.
2. Library-size normalization (Wolf, Angerer, and Theis 2018): `scanpy.pp.normalize_total()` with default parameters (Norman, Dixit, Papalexi) or `target_sum=1e4` (Replogle, Adamson).
3. Log transformation: `scanpy.pp.log1p()`.
4. Highly variable gene selection: `scanpy.pp.highly_variable_genes(n_top_genes=2000, subset=True)`.
5. PCA: `scanpy.tl.pca(n_comps=50)`.

All downstream stability and magnitude computations were performed on the 50-dimensional PCA embedding.

Control group identification

Control group assignment used a multi-stage matching protocol to accommodate heterogeneous labeling conventions across datasets:

1. **Exact match** (case-insensitive): labels matching "control", "ctrl", "non-targeting", "NT", "unperturbed".
2. **Delimiter-aware regex**: for short tokens (e.g., "NT" in "NTg1"), split on common delimiters and match components.
3. **Substring matching**: for longer keywords embedded in complex labels.

Dataset-specific handling:

- **Replogle 2022:** Labels containing "non-targeting" or beginning with "chr" were assigned to control. Labels containing "pos_control" were removed. Gene names were extracted by splitting on underscore and taking the first token.
- **Papalexi 2021:** The `gene_target` column was copied from MuData global metadata. Non-targeting guides (NTg1 through NTg7) were pooled into a single NT control group (2,386 cells). Note: the pertpy loader for Papalexi is incompatible with current mudata versions due to hashed MuData keys; results for this dataset are from a prior pipeline run with identical preprocessing parameters.
- **Norman 2019:** The `perturbation_name` column was used directly. Cells labeled "control" served as the control group.

Perturbation stability and effect magnitude

For a perturbation p applied to n_p cells, let $\mathbf{x}_i \in \mathbb{R}^{50}$ denote the PCA coordinates of perturbed cell i and $\boldsymbol{\mu}_{\text{ctrl}}$ the centroid of all control cells in the same dataset. The shift vector for cell i is:

$$\mathbf{d}_i = \mathbf{x}_i - \boldsymbol{\mu}_{\text{ctrl}} \tag{2}$$

The mean perturbation direction is:

$$\bar{\mathbf{d}} = \frac{1}{n_p} \sum_{i=1}^{n_p} \mathbf{d}_i \tag{3}$$

Perturbation stability (Shesha) is defined as:

$$S_p = \frac{1}{n_p} \sum_{i=1}^{n_p} \frac{\mathbf{d}_i \cdot \bar{\mathbf{d}}}{\|\mathbf{d}_i\| \, \|\bar{\mathbf{d}}\|} \tag{4}$$

Effect magnitude is:

$$M_p = \|\bar{\mathbf{d}}\| \tag{5}$$

Discordance is the standardized residual from the ordinary least-squares regression of S_p on M_p:

$$D_p = z(M_p) - z(S_p) \tag{6}$$

where $z(\cdot)$ denotes within-dataset z-score normalization. Positive discordance indicates that a perturbation has lower stability than predicted by its magnitude (below the regression line); negative discordance indicates higher stability than predicted (above the line).

The perturbation stability metric adapts the principle of geometric self-consistency from the Shesha representational stability framework (Raju 2026a,b). In the general framework, Shesha-FS measures split-half RDM consistency across representations. Here, S_p measures perturbation coherence directly within a single representation, specializing the principle to perturbation biology.

Minimum cell count filtering

Perturbations with fewer than 50 cells (10 for Dixit) were excluded to ensure stable estimates. The Adamson dataset retains all 8 perturbations despite wide bootstrap confidence intervals, which honestly reflect the limited statistical power.

Robustness Analyses

Distance metric robustness

All methods produce comparable or stronger correlations relative to the standard Euclidean metric (Table 3), confirming that the magnitude-stability relationship is robust to the choice of distance metric. Whitened (Mahalanobis-scaled) and k-NN matched methods produce equal or higher correlations across all four datasets, consistent with these methods reducing noise from batch effects and control group heterogeneity.

PCA dimensionality ablation

Stability was recomputed at PCA dimensionalities of 10, 20, 30, 50, and 100 components. Spearman ρ increased monotonically with dimensionality: Norman $\rho = 0.949$ (10 PCs) to 0.969 (100 PCs); Dixit $\rho = 0.781$ (10 PCs) to 0.869 (100 PCs). The default of 50 components ($\rho = 0.959$ for Norman, 0.844 for Dixit) lies in the upper portion of the range, and the relationship is robust across all tested dimensionalities.

Seed reproducibility

Cross-seed Spearman correlation of perturbation-level stability rankings across 15 random seeds (320, 1991, 9, 7258, 7, 2222, 724, 3, 12, 108, 18, 11, 1754, 411, 103) was near-perfect: Norman mean $r = 0.99997$ (range $[0.9999, 1.0000]$, 105 pairwise comparisons); Dixit mean $r = 0.99963$ (range $[0.9993, 0.9999]$). The magnitude-stability Spearman ρ varied by less than 0.001 across seeds for both datasets (Norman: 0.959–0.960; Dixit: 0.844–0.844). All reported results use seed 320.

Leave-one-out influence

The maximum $|\Delta\rho|$ upon removing any single perturbation was 0.0019 for Norman (most influential: HES7; LOO ρ range $[0.959, 0.961]$) and 0.0106 for Dixit (most influential: INTERGENIC1144056+INTERGENIC1216445; LOO ρ range $[0.840, 0.854]$), indicating that no individual perturbation drives the overall correlation.

Theoretical null model

Under a null model where individual cell shift vectors are drawn uniformly on the unit hypersphere (50 dimensions), the expected stability is $S_p \approx 0$ with variance inversely proportional to n_p. Observed stability values (0.05–0.85) far exceed this null, confirming that the coherence signal is biological rather than statistical.

Mixed-Effects Model

To assess cross-dataset generalization and identify confounds, we fit a linear mixed-effects model:

$$S_p = \beta_0 + \beta_1 \cdot M_p + \beta_2 \cdot \text{spread}_p + \beta_3 \cdot n_p + u_{\text{dataset}} + \varepsilon_p \tag{7}$$

where M_p is effect magnitude, spread_p is within-perturbation expression variance, n_p is cell count, u_{dataset} is a dataset-level random intercept, and ε_p is residual error. All predictors were z-scored within each dataset prior to fitting.

Magnitude accounts for approximately 11 times more variance in stability than sample size ($|\beta_1/\beta_3| = 11.2$), confirming that the magnitude-stability relationship is not driven by differential cell sampling. The near-zero dataset random-effect variance confirms that the relationship generalizes across datasets rather than being driven by any single screen. Full results in Table 4.

Extended Stress Marker Analysis

Partial correlations

Table 15 reports raw and partial Spearman correlations between perturbation stability and four stress/UPR markers across three datasets, controlling for effect magnitude.

Quadrant depletion tests

Perturbations were split at median stability and median stress marker expression. The high-stability/high-stress (HH) quadrant count was compared to expectation under independence using a one-sided binomial test. Full quadrant counts for all four markers across all three datasets are provided in Table 14.

Modality analysis

The DDIT3 sign heterogeneity (positive raw ρ in Norman and Replogle, negative in Dixit) partly reflects CRISPRa vs CRISPRi biology. However, within CRISPRi alone, the sign also differs between Dixit (BMDCs, raw $\rho = -0.365$) and Replogle (K562, raw $\rho = +0.382$), indicating that cell type and experimental design contribute to the heterogeneity. After magnitude control, all three DDIT3 partial correlations are negative, but only Replogle survives with a small effect.

HSPA5 is the most consistent marker across modalities: negative partial correlations in both CRISPRi datasets (Dixit and Replogle), null in CRISPRa (Norman). Full modality-stratified results are provided in Tables 15.

Extended Replogle Analysis

The Replogle 2022 genome-scale CRISPRi screen provides independent validation at unprecedented scale (n = 1,832 perturbations after filtering for $\geq$ 50 cells; 310,385 total cells). The Replogle discordance scatter (*SI Appendix*, Fig. S1) shows the same pattern observed in Norman: pleiotropic regulators (GATA1, CHMP3, AQR) cluster below the regression line, while narrowly-acting factors (LSG1, ISG20L2, KRI1) cluster above. See Table 7.

scGPT Validation Protocol

Model and embedding

Cell embeddings were generated using the scGPT "Whole Human" pretrained checkpoint (`scGPT_human`), downloaded from the official repository (`https://github.com/bowang-lab/scGPT`). The checkpoint contains three files: `best_model.pt`, `vocab.json`, and `args.json`.

Input data consisted of raw counts (the `counts` layer was extracted from each AnnData object; data were not log-normalized prior to embedding). Embeddings were generated using `embed_data()` with `gene_col="index"`, `batch_size=64`, and `use_fast_transformer=False`.

Reproducibility settings

All scGPT computations used deterministic mode:

- Python random seed: 320
- NumPy random seed: 320
- PyTorch manual seed: 320 (CPU and CUDA)
- `torch.backends.cudnn.deterministic = True`
- `torch.backends.cudnn.benchmark = False`

Stability and magnitude computation

Stability and magnitude were computed from scGPT embeddings using `shesha.bio.compute_stability()` and `shesha.bio.compute` with `perturbation_key='perturbation_name'`, `control_label='control'`, and `metric='cosine'` (stability) or `metric='euclide` (magnitude). Bootstrap confidence intervals: 10,000 resamples, seed 320, percentile method.

Datasets

scGPT (Cui et al. 2024) analysis was performed on Norman 2019, Dixit 2016, and Replogle 2022. Adamson 2016 ($n = 8$) was included in the output CSV but omitted from main-text figures due to limited power. Papalexi 2021 was excluded due to pertpy loader incompatibility with current mudata versions.

Results

The dataset rank order is preserved (Norman > Replogle > Dixit). scGPT correlations are consistently slightly lower, with the largest drop in Replogle ($\Delta\rho = -0.119$), consistent with the nonlinear embedding resolving manifold structure that PCA collapses. See Table 13

Combinatorial Analysis

The Norman 2019 dataset contains $n = 105$ single-gene and $n = 131$ combinatorial (two-gene) perturbations. Combinatorial perturbations showed significantly higher stability (mean $S_p = 0.460$ vs 0.306; Mann-Whitney $U = 2{,}637$, $p = 4.1 \times 10^{-16}$).

The magnitude-stability relationship held within both categories, with regression slopes of 0.089 (single-gene, $R = 0.925$, $\rho = 0.973$) and 0.058 (combinatorial, $R = 0.874$, $\rho = 0.919$), confirming that the higher stability of combinatorial perturbations is not simply a consequence of their larger effect magnitudes.

Within the combinatorial set, perturbations involving lineage-specific factors (e.g., KLF1+SET, KLF1+TGFBR2) showed higher stability than those involving pleiotropic factors (e.g., CEBPA+JUN, CEBPA+CEBPB), consistent with the discordance pattern observed in the single-gene analysis.

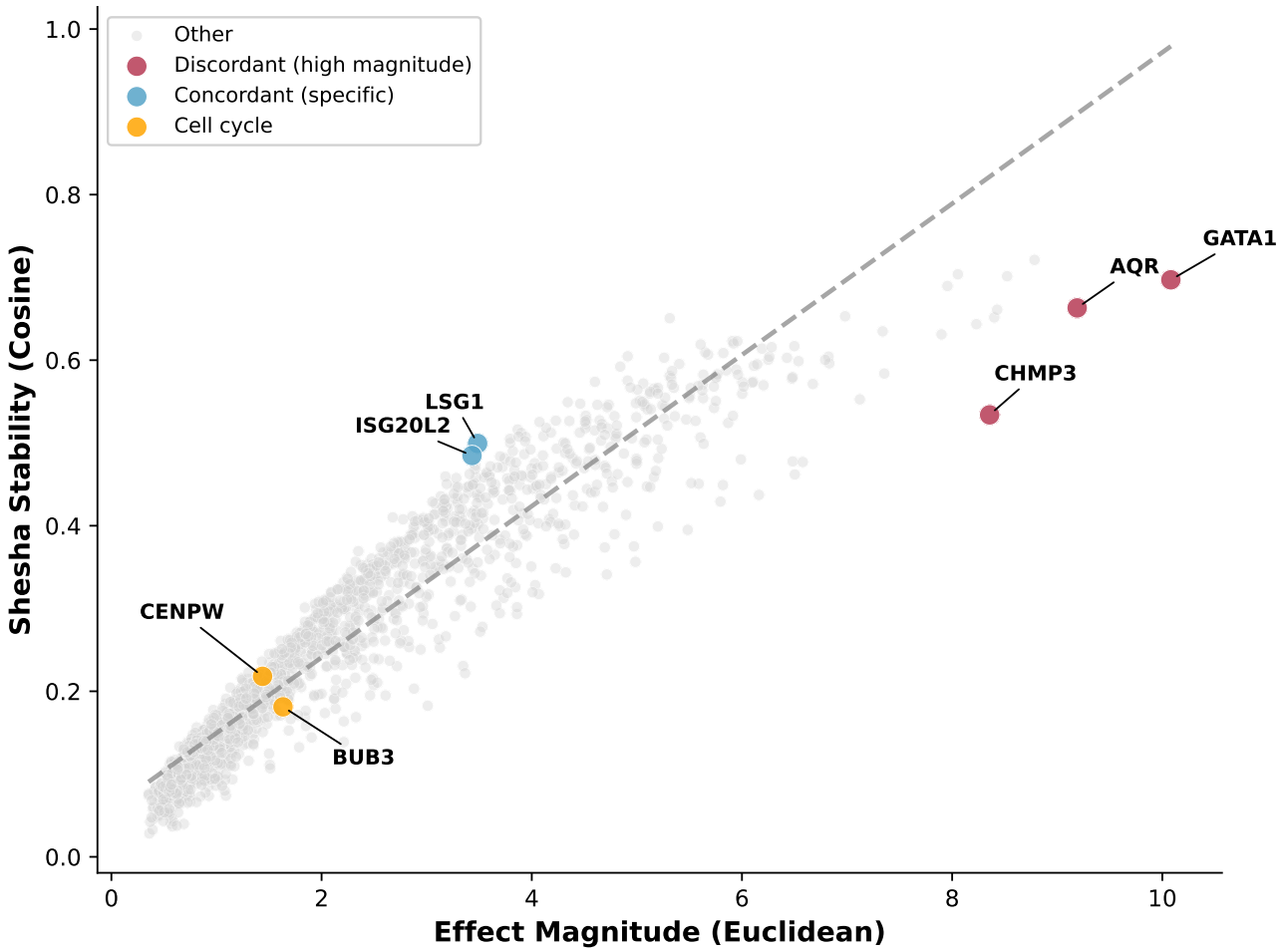

Figure 6: **Extended discordance analysis in the Replogle 2022 genome-scale CRISPRi screen.** Effect magnitude (Euclidean, x-axis) versus Shesha stability (cosine, y-axis) for n = 1,832 perturbations in Replogle et al. (2022) K562 cells. Dashed line shows linear fit. Points are categorized by biological function: *Discordant* (red): high magnitude but low stability relative to regression, including GATA1 (master regulator of erythroid/megakaryocytic differentiation, discordance = 2.15), CHMP3 (ESCRT-III complex, discordance = 2.14), and ACTB (β-actin, essential cytoskeletal gene). *Concordant* (blue): high stability relative to magnitude, including LSG1 (ribosome biogenesis factor) and related factors with narrow downstream programs. *Cell cycle* (orange): BUB3 (spindle assembly checkpoint) and CENPW (centromere protein) show low stability independent of cell cycle arrest, demonstrating that geometric incoherence is not merely a proxy for proliferation block. The discordance pattern independently validates the Geometric Tax observed in Norman CRISPRa: pleiotropic master regulators produce large but incoherent effects, while pathway-specific factors produce coherent shifts.

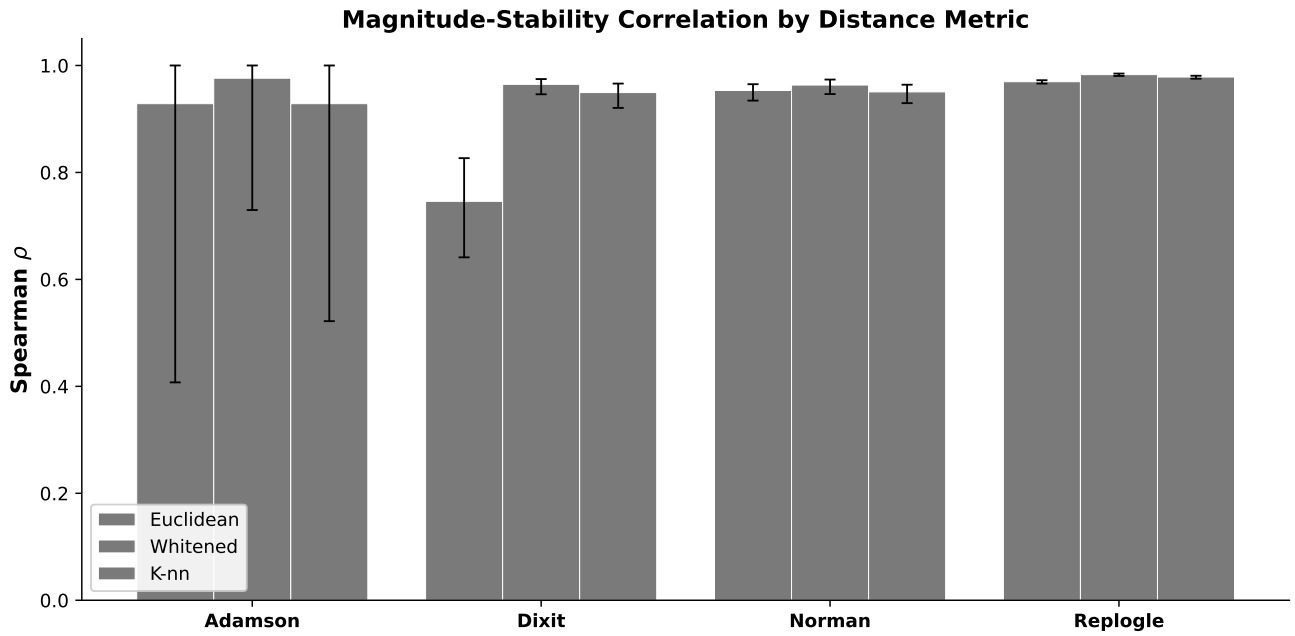

Figure 7: **Magnitude-stability correlation is robust across distance metrics.** Bar chart showing Spearman correlations with 95% bootstrap CIs (error bars) for three distance computation methods: Euclidean (standard L_2 in PCA space), Whitened (Mahalanobis-scaled coordinates), and k-NN (local control centroids). All methods achieve strong correlations ($\rho > 0.67$) across all datasets. Whitening substantially improves the Dixit correlation from $\rho = 0.75$ to $\rho = 0.97$, suggesting residual covariance structure in PCA space attenuates the relationship in that dataset.

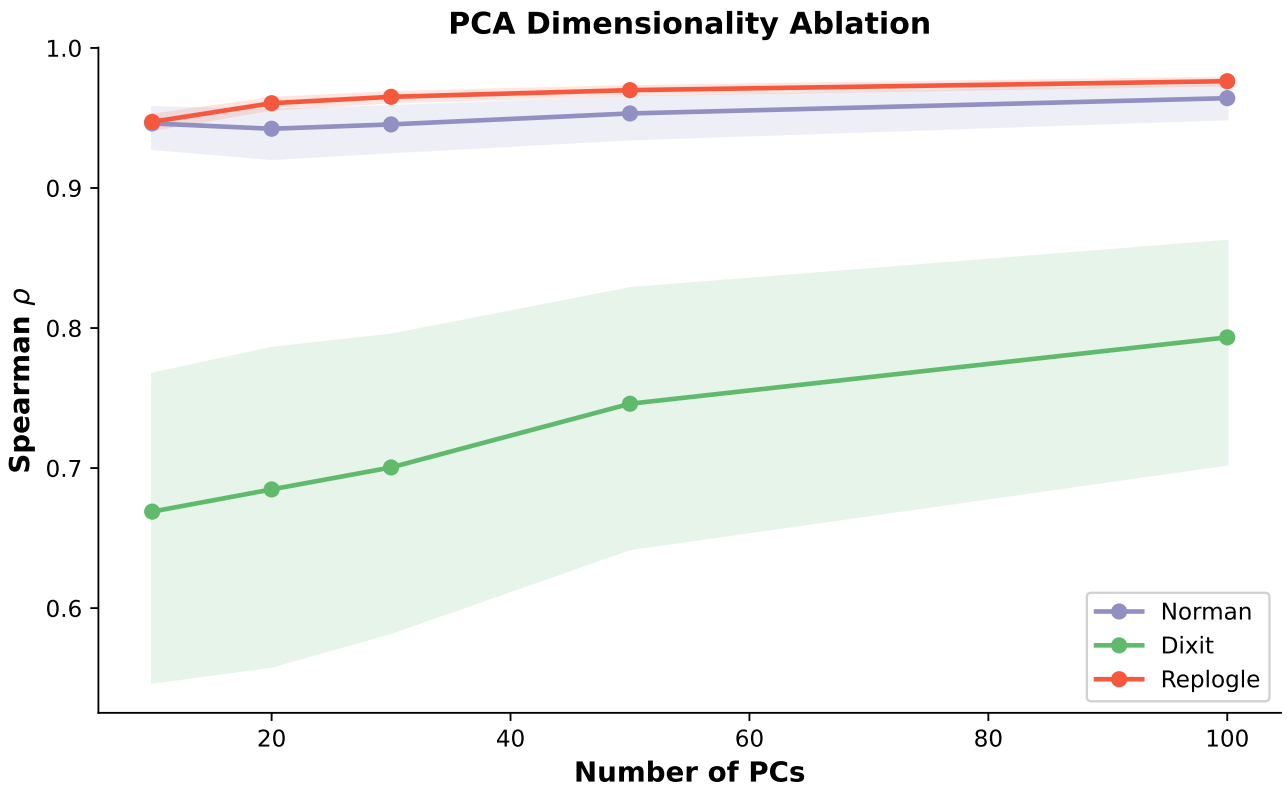

Figure 8: **PCA dimensionality ablation.** Magnitude-stability Spearman ρ as a function of the number of principal components retained (10, 20, 30, 50, 100). Shaded regions indicate 95% bootstrap CIs (10,000 iterations). Norman 2019 shows stable correlations ($\rho = 0.94$–0.96) with overlapping CIs across all settings. Dixit 2016 shows modest improvement with more PCs ($\rho = 0.67$ to 0.79), suggesting higher-dimensional structure contributes to the relationship in that dataset. Replogle 2022 shows consistently high correlations ($\rho = 0.95$–0.98). Rank-order perturbation consistency is high across all settings (Norman: $r = 0.98 \pm 0.02$; Dixit: $r = 0.96 \pm 0.04$), confirming that the choice of 50 PCs does not drive the results.

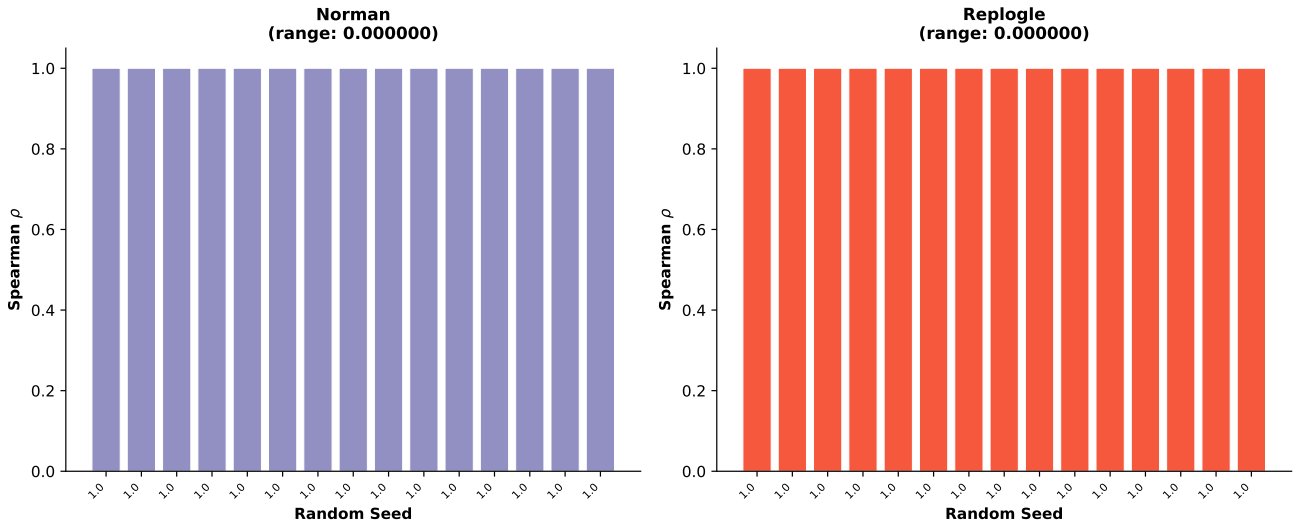

Figure 9: **Random seed reproducibility.** Magnitude-stability Spearman ρ recomputed across 15 different random seeds ($\{3, 7, 9, 11, 12, 18, 103, 108, 320, 411, 724, 1754, 1991, 2222, 7258\}$) for Norman 2019 and Replogle 2022. All correlations are identical to machine precision (cross-seed $r = 1.000$), confirming that stochastic elements in the preprocessing pipeline (PCA initialization) have no effect on the final results.

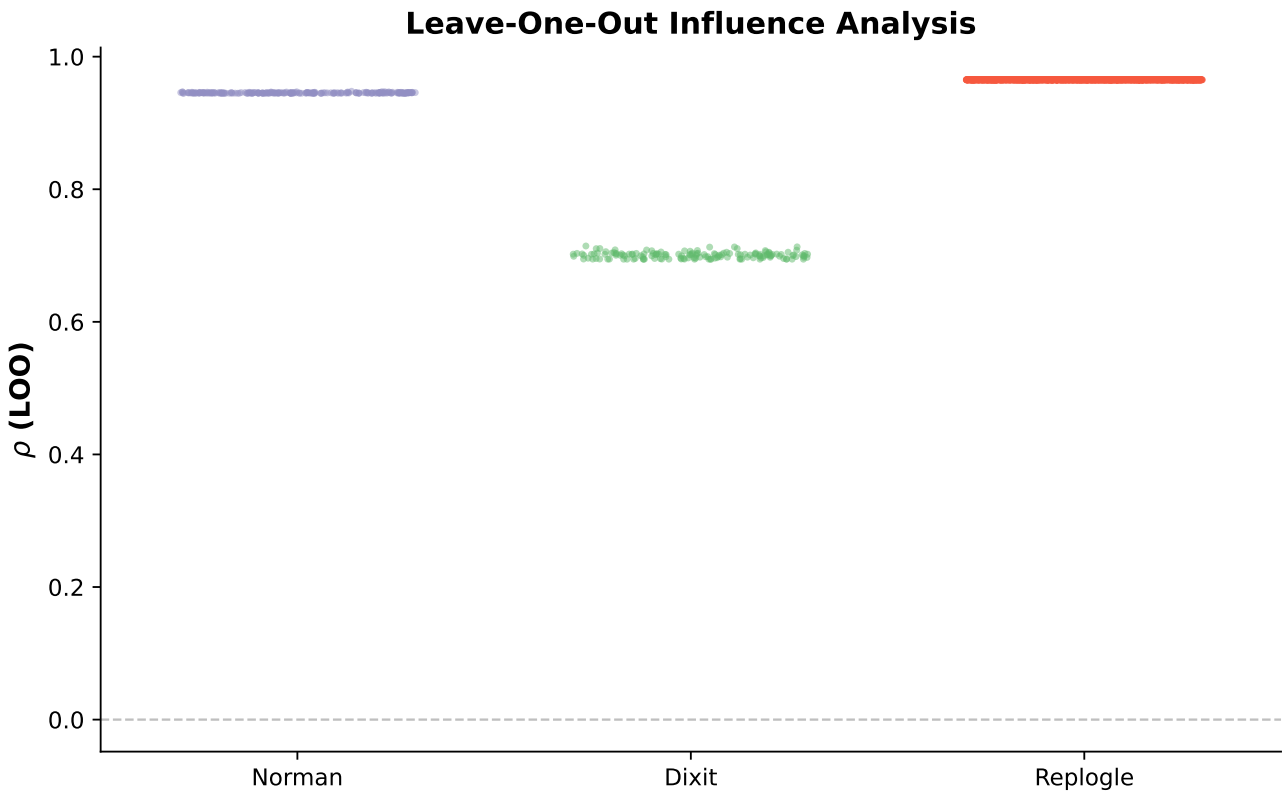

Figure 10: **Leave-one-out influence analysis.** Distribution of $\Delta\rho$ values when each perturbation is removed in turn. The LOO range is narrow for all datasets: removing any single perturbation changes the correlation by at most $\Delta\rho = 0.002$ (Norman), $\Delta\rho = 0.014$ (Dixit), or $\Delta\rho < 0.001$ (Replogle). Most influential perturbations: BAK1 (most helpful, Norman), HES7 (most harmful, Norman), ELK1 (most helpful, Dixit), CREB1+E2F4+ELF1 (most harmful, Dixit).

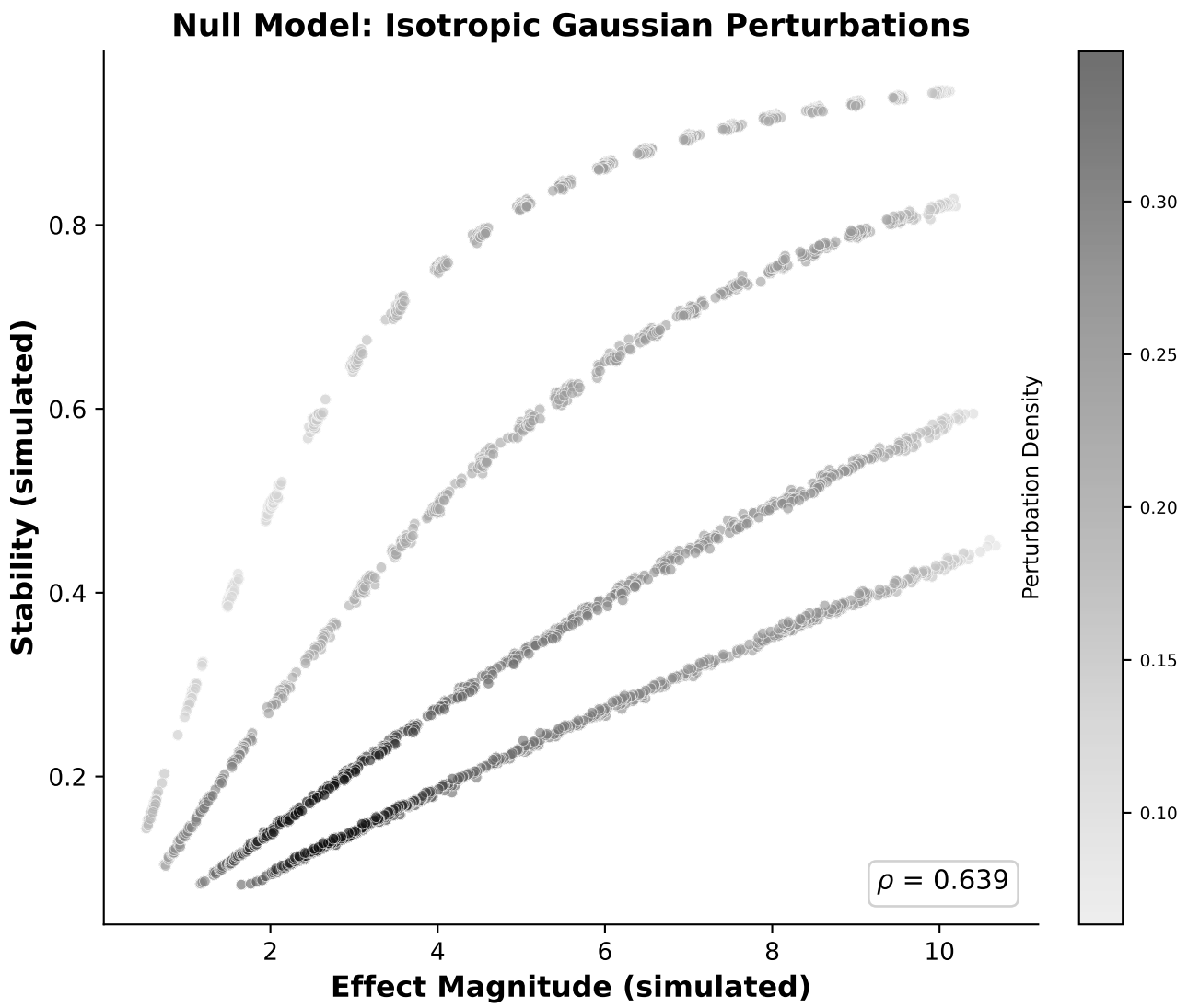

Figure 11: **Theoretical null model under isotropic Gaussian perturbations.** Magnitude (x-axis) versus stability (y-axis) for 2,000 simulated perturbations ($d = 50$ dimensions, $\sigma \in \{0.5, 1.0, 2.0, 3.0\}$, 500 simulations per condition). Under the null model, stability is almost perfectly predicted by SNR ($\rho = 0.999$), with a partial correlation of $\rho_{\text{partial}} = 0.292$ after controlling for SNR. The heterogeneity observed in real data (Norman $\rho_{\text{partial}} = -0.859$, Dixit $\rho_{\text{partial}} = +0.627$) far exceeds this null prediction, confirming that biological factors beyond simple SNR confounding drive the magnitude-stability relationship.

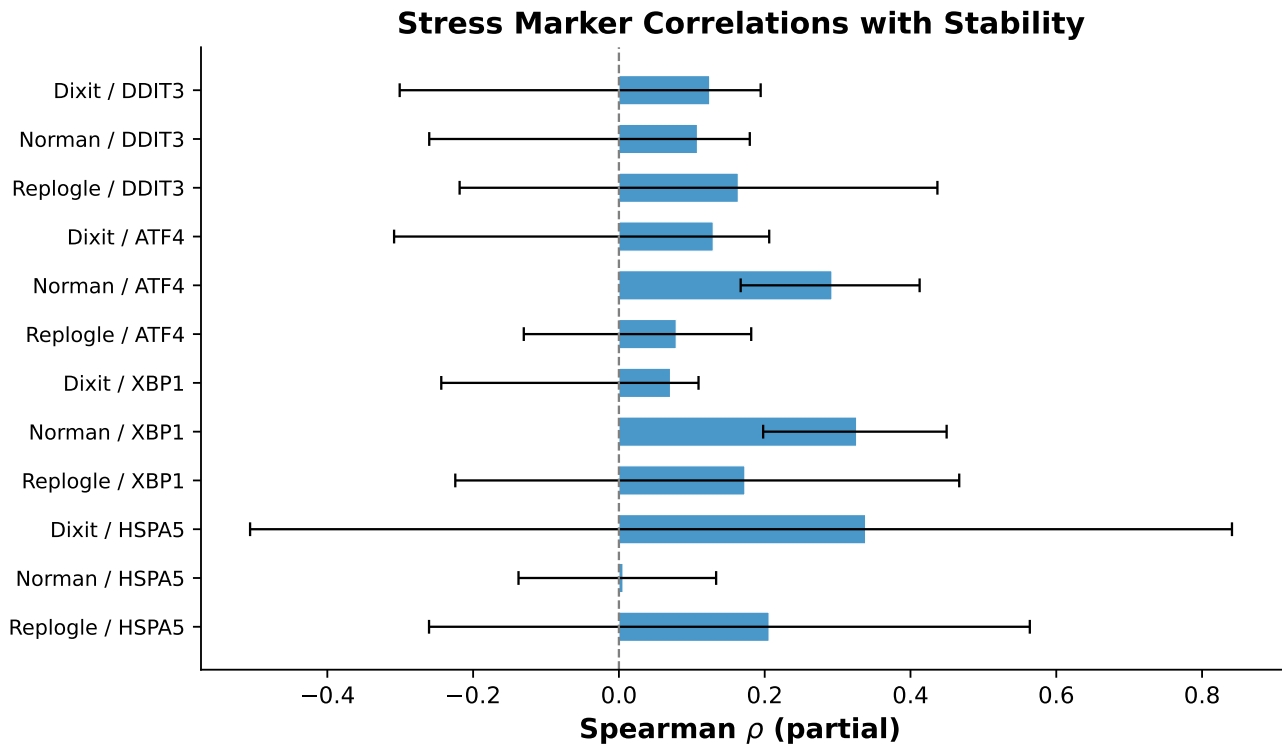

Figure 12: **Stress marker correlations with geometric stability.** Forest plot of Spearman correlations between perturbation stability (S_p) and mean expression of four canonical stress response markers (DDIT3, ATF4, XBP1, HSPA5) across three datasets (Norman, Dixit, Replogle). Bars extend to 95% bootstrap CIs. Significant associations ($p < 0.001$) in bold. HSPA5 shows the most consistent negative association across datasets ($\rho = -0.31$ to -0.40 in Dixit and Replogle), while DDIT3 shows sign heterogeneity between CRISPRa and CRISPRi modalities, reflecting the directional effect of activation versus interference on stress pathway engagement.

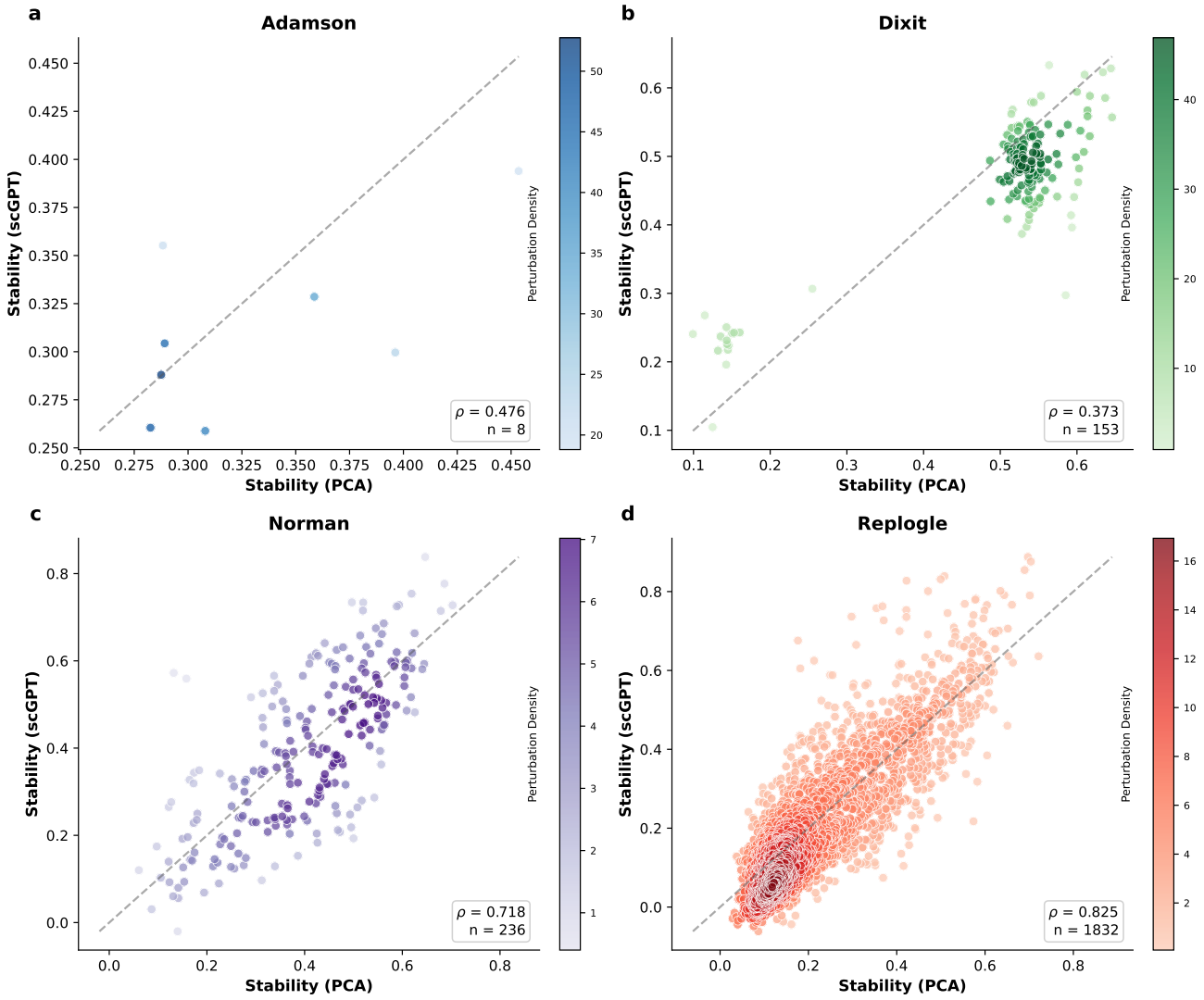

Figure 13: **Per-perturbation concordance between PCA and scGPT stability estimates.** Each panel shows PCA-derived stability (x-axis) versus scGPT-derived stability (y-axis) for shared perturbations, with identity line (dashed). Point color indicates local perturbation density. Spearman ρ of paired values annotated. High concordance confirms that stability rankings are preserved across linear and nonlinear embedding spaces, and that the magnitude-stability relationship is a property of biological state space rather than an artifact of the dimensionality reduction method.

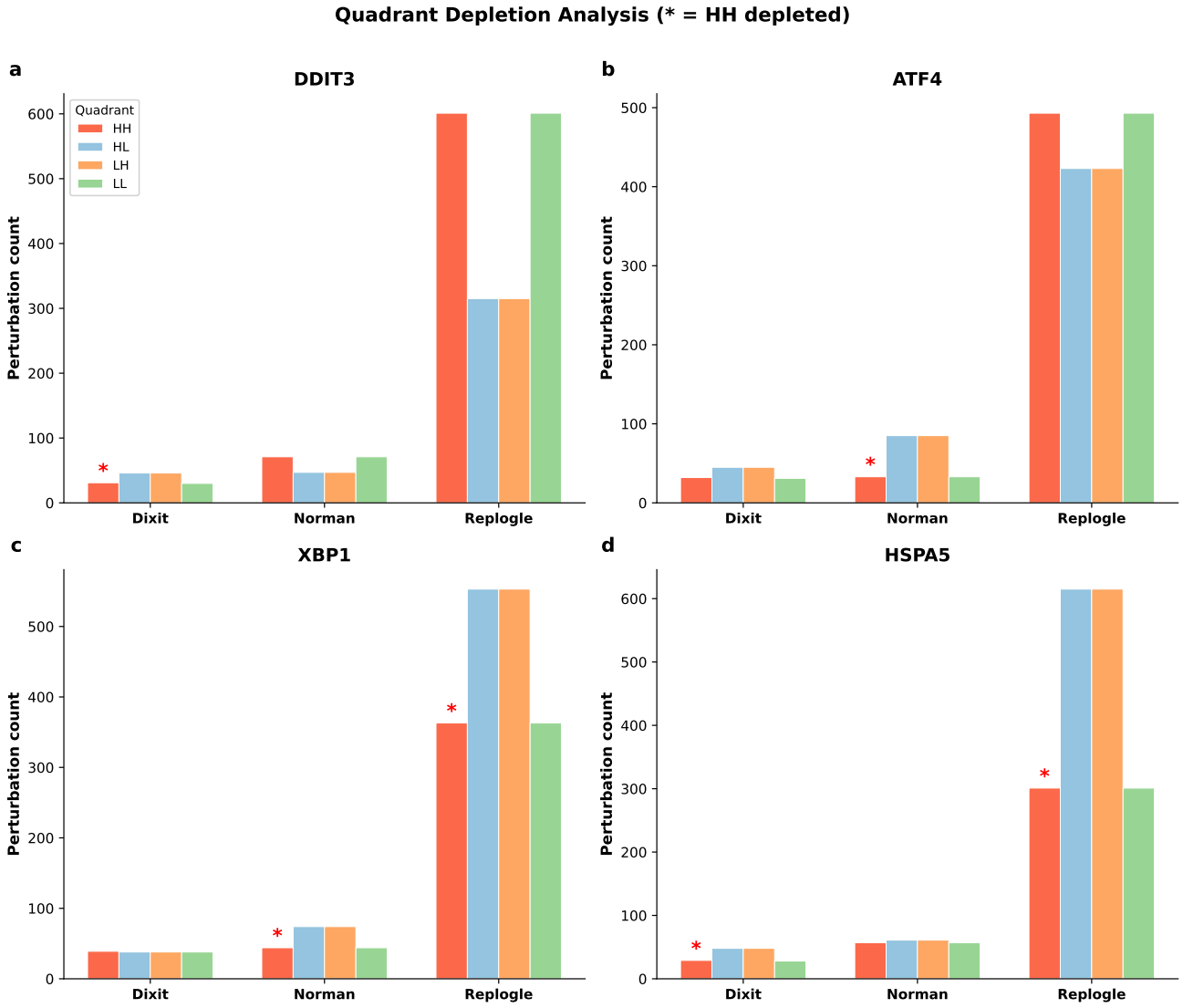

Figure 14: **Quadrant depletion analysis of stability versus stress.** Perturbations split at median stability and median stress expression (dashed lines). Quadrant counts annotated. The high-stability/high-stress (HH) quadrant is systematically depleted across multiple stress markers and datasets (Fisher's exact test; Table 14), supporting the interpretation that geometric coherence is a prerequisite for cellular homeostasis. Perturbations producing incoherent cellular responses (low stability) are more likely to induce elevated stress signatures.

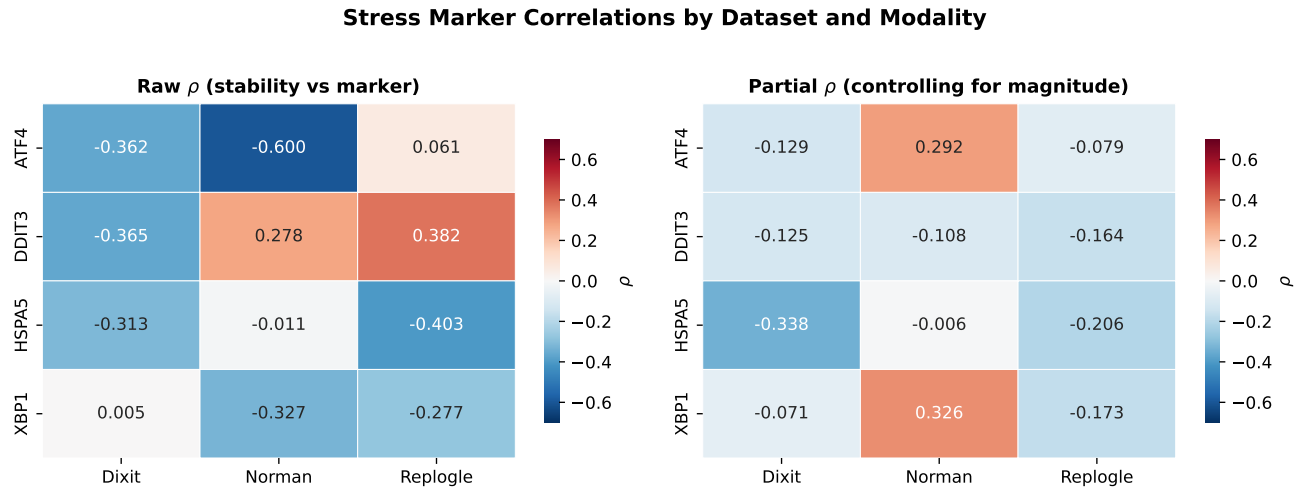

Figure 15: **Stress marker correlations by dataset and modality.** Heatmap of Spearman correlations between geometric stability and four stress markers (DDIT3, ATF4, XBP1, HSPA5) across three datasets. Color scale: blue (positive) to red (negative), centered at zero. The heterogeneity across markers and datasets reflects differences in baseline stress levels, perturbation modality (CRISPRa versus CRISPRi), and the specific stress pathway engaged by each class of perturbation.

Table 2: Dataset overview.

Dataset	Modality	Cell type	Perturbations	Total cells	Median cells/pert	Control label
Norman 2019 (Norman et al. 2019)	CRISPRa	K562	236	111,255	352	control
Adamson 2016 (Adamson et al. 2016)	CRISPRi	K562	8	5,752	560	control
Dixit 2016 (Dixit et al. 2016)	CRISPRi	BMDCs	153	99,722	75	control
Papalexi 2021(Papalexi et al. 2021)	Pooled	THP-1	25	18,343	662	NT (pooled)
Replogle 2022 (Replogle et al. 2022)	CRISPRi	K562	1,832	310,385	132	non-targeting

Table 3: Spearman ρ between magnitude and stability under different distance metrics.

Dataset	Euclidean	Whitened	k-NN
Adamson 2016	0.929	0.976	0.929
Dixit 2016	0.746	0.965	0.949
Norman 2019	0.953	0.963	0.951
Replogle 2022	0.970	0.983	0.978

Table 4: Mixed-effects model results.

Parameter	β	95% CI
Magnitude	0.168	[0.166, 0.170]
Spread	−0.122	[−0.128, −0.116]
Sample size	−0.015	[−0.017, −0.013]
Dataset variance	≈ 0	—

Table 5: Magnitude-stability Spearman correlations with 95% bootstrap CIs (10,000 iterations).

Dataset	n	ρ	95% CI	p
Norman	236	0.953	[0.934, 0.965]	$< 10^{-100}$
Adamson	8	0.929	[0.407, 1.000]	$< 10^{-4}$
Dixit	153	0.746	[0.641, 0.827]	$< 10^{-28}$
Replogle	1832	0.970	[0.966, 0.972]	$< 10^{-100}$
Papalexi	25	0.985	[0.939, 0.997]	$< 10^{-19}$

Table 6: Magnitude-stability correlation across distance metrics.

Dataset	$\rho_{\text{Euclidean}}$	ρ_{Whitened}	$\rho_{\text{K-nn}}$
Adamson	0.929 [0.407, 1.000]	0.976 [0.730, 1.000]	0.929 [0.522, 1.000]
Dixit	0.746 [0.641, 0.827]	0.965 [0.946, 0.975]	0.949 [0.921, 0.966]
Norman	0.953 [0.934, 0.965]	0.963 [0.947, 0.974]	0.951 [0.930, 0.964]
Replogle	0.970 [0.966, 0.972]	0.983 [0.980, 0.985]	0.978 [0.975, 0.981]

Table 7: Top 10 discordant and top 5 concordant perturbations in Replogle 2022 (PCA space).

Gene	Magnitude	Stability	Discordance	*n* cells	Function
Most discordant (high magnitude, low stability relative to fit)					
GATA1	10.081	0.697	2.15	108	Erythroid/megakaryocytic TF
CHMP3	8.357	0.534	2.14	130	ESCRT-III complex
AQR	9.189	0.663	1.80	90	RNA helicase
SLU7	6.503	0.462	1.42	225	Splicing factor
VPS28	6.579	0.477	1.37	59	ESCRT-I complex
PHB	8.401	0.652	1.37	256	Prohibitin (mitochondrial)
CNOT1	6.162	0.437	1.37	269	CCR4-NOT deadenylase
INTS2	8.428	0.661	1.32	354	Integrator complex
SUPT6H	8.230	0.643	1.31	288	Transcription elongation
HSPA5	6.483	0.478	1.30	438	ER chaperone (BiP/GRP78)
Most concordant (high stability relative to magnitude)					
LSG1	3.485	0.499	−0.79	116	Ribosome biogenesis
ISG20L2	3.431	0.485	−0.72	73	RNA exonuclease
KRI1	3.178	0.460	−0.72	52	Ribosome biogenesis
LSM10	2.677	0.410	−0.71	135	snRNP assembly
POLD3	3.866	0.516	−0.66	68	DNA polymerase delta

Table 8: PCA dimensionality ablation. Magnitude-stability correlations (ρ) with 95% bootstrap CIs (10,000 iterations) across varying numbers of principal components.

Dataset	PCs	ρ	95% CI	p
Norman	10	0.946	[0.928, 0.958]	$< 10^{-116}$
	20	0.942	[0.921, 0.956]	$< 10^{-113}$
	30	0.945	[0.926, 0.958]	$< 10^{-116}$
	50	0.953	[0.935, 0.965]	$< 10^{-123}$
	100	0.964	[0.949, 0.973]	$< 10^{-137}$
Dixit	10	0.669	[0.547, 0.767]	$< 10^{-21}$
	20	0.685	[0.558, 0.786]	$< 10^{-22}$
	30	0.700	[0.582, 0.795]	$< 10^{-24}$
	50	0.746	[0.642, 0.828]	$< 10^{-28}$
	100	0.793	[0.703, 0.862]	$< 10^{-34}$
Replogle	10	0.947	[0.941, 0.952]	$< 10^{-300}$
	20	0.961	[0.956, 0.964]	$< 10^{-300}$
	30	0.965	[0.961, 0.968]	$< 10^{-300}$
	50	0.970	[0.966, 0.973]	$< 10^{-300}$
	100	0.976	[0.973, 0.979]	$< 10^{-300}$

Table 9: Random seed reproducibility. Stability recomputed using 15 different random seeds per dataset. All correlations are identical to machine precision (cross-seed $r = 1.000$), confirming no stochastic dependence in the analytic pipeline.

Dataset	n	Seeds	ρ (mean ± std)	95% CI	Cross-seed r
Norman	236	15	0.945403 ± 0.000000	[0.925, 0.959]	1.000
Replogle	1832	15	0.965106 ± 0.000000	[0.961, 0.969]	1.000

Table 10: Leave-one-out influence analysis. Removing any single perturbation changes the correlation by at most $\Delta\rho = 0.014$ (Dixit) or $\Delta\rho = 0.002$ (Norman), confirming no individual perturbation drives the observed relationship.

Dataset	n	Full ρ	LOO range	Most helpful ($\Delta\rho$)	Most harmful ($\Delta\rho$)
Norman	236	0.945	[0.945, 0.948]	BAK1 (+0.0007)	HES7 (−0.0024)
Dixit	153	0.700	[0.695, 0.714]	ELK1 (+0.0060)	CREB1+E2F4+ELF1 (−0.0139)
Replogle	1832	0.965	[0.965, 0.965]	EIF2S1 (+0.0001)	CRNKL1 (−0.0004)

Table 11: Partial correlations between stability and magnitude, controlling for intrinsic spread and sample size. The heterogeneity across datasets (Norman $\rho_{\text{partial}} = -0.86$, Dixit $\rho_{\text{partial}} = +0.63$) exceeds what the isotropic null model predicts, indicating biological factors beyond SNR confounding.

Dataset	ρ_{partial}	95% CI	p
Norman	-0.859	$[-0.905, -0.781]$	$< 10^{-70}$
Dixit	0.627	$[0.482, 0.728]$	$< 10^{-18}$
Replogle	-0.789	$[-0.812, -0.765]$	$< 10^{-300}$
Pooled	-0.102	$[-0.156, -0.049]$	$< 10^{-6}$

Table 12: Stress marker correlations with geometric stability. Spearman correlations between perturbation stability (S_p) and mean expression of canonical stress markers. 95% bootstrap CIs (10,000 iterations). Significant results ($p < 0.001$) in **bold**.

Dataset	**DDIT3**		**ATF4**		**XBP1**		**HSPA5**	
	ρ	CI	ρ	CI	ρ	CI	ρ	CI
Norman	**0.278**	[0.16, 0.39]	**−0.600**	[−0.68, −0.51]	**−0.327**	[−0.44, −0.20]	−0.011	[−0.14, 0.12]
Dixit	**−0.365**	[−0.52, −0.20]	**−0.362**	[−0.52, −0.19]	0.005	[−0.15, 0.16]	**−0.313**	[−0.47, −0.15]
Replogle	**0.382**	[0.34, 0.43]	0.061	[0.01, 0.11]	**−0.277**	[−0.33, −0.23]	**−0.403**	[−0.45, −0.36]

Table 13: PCA vs scGPT magnitude-stability correlations.

Dataset	PCA ρ	scGPT ρ	scGPT 95% CI	$\Delta\rho$
Norman 2019	0.953	0.935	[0.911, 0.951]	−0.018
Dixit 2016	0.746	0.712	[0.585, 0.818]	−0.034
Replogle 2022	0.970	0.851	[0.836, 0.865]	−0.119

Table 14: Quadrant depletion tests. Perturbations split at median stability and median stress expression. The high-stability/high-stress (HH) quadrant is tested for depletion via Fisher's exact test. Significant depletions (**bold**) suggest geometric coherence is a prerequisite for cellular homeostasis.

Dataset	Marker	n	HH	HL	LH	LL	Fisher p	Depleted?
Dixit	**DDIT3**	**153**	**31**	**46**	**46**	**30**	**0.015**	**Yes**
Norman	DDIT3	236	71	47	47	71	0.003	No
Replogle	DDIT3	1832	601	315	315	601	$< 10^{-41}$	No
Dixit	ATF4	153	32	45	45	31	0.036	No
Norman	**ATF4**	**236**	**33**	**85**	**85**	**33**	$< 10^{-11}$	**Yes**
Replogle	ATF4	1832	493	423	423	493	0.001	No
Dixit	XBP1	153	39	38	38	38	1.000	No
Norman	**XBP1**	**236**	**44**	**74**	**74**	**44**	$< 10^{-4}$	**Yes**
Replogle	**XBP1**	**1832**	**363**	**553**	**553**	**363**	$< 10^{-19}$	**Yes**
Dixit	**HSPA5**	**153**	**29**	**48**	**48**	**28**	**0.002**	**Yes**
Norman	HSPA5	236	57	61	61	57	0.696	No
Replogle	**HSPA5**	**1832**	**301**	**615**	**615**	**301**	$< 10^{-49}$	**Yes**

Table 15: Raw and partial correlations between stability and stress markers. Partial correlations control for effect magnitude. Bold: survives magnitude control (CI excludes zero and effect size ≥ small).

Dataset	Marker	n	Raw ρ	Partial ρ	Partial CI	p	Survives
Dixit	DDIT3	153	−0.365	−0.125	[−0.301, 0.055]	0.125	No
Norman	DDIT3	236	+0.278	−0.108	[−0.260, 0.036]	0.099	No
Replogle	DDIT3	1,832	+0.382	−0.164	[−0.219, −0.109]	1.7×10^{-12}	**Yes**
Dixit	ATF4	153	−0.362	−0.129	[−0.308, 0.053]	0.111	No
Norman	ATF4	236	−0.600	+0.292	[0.167, 0.413]	5.0×10^{-6}	**Yes**
Replogle	ATF4	1,832	+0.061	−0.079	[−0.130, −0.024]	0.001	No*
Dixit	XBP1	153	+0.005	−0.071	[−0.244, 0.109]	0.384	No
Norman	XBP1	236	−0.327	+0.326	[0.198, 0.450]	3.0×10^{-7}	**Yes**
Replogle	XBP1	1,832	−0.277	−0.173	[−0.224, −0.121]	9.4×10^{-14}	**Yes**
Dixit	HSPA5	153	−0.313	**−0.338**	[−0.506, −0.164]	1.9×10^{-5}	**Yes**
Norman	HSPA5	236	−0.011	−0.006	[−0.138, 0.133]	0.932	No
Replogle	HSPA5	1,832	−0.403	**−0.206**	[−0.260, −0.152]	5.2×10^{-19}	**Yes**

*Replogle ATF4: CI excludes zero due to large n, but effect size is negligible ($|\rho| = 0.079$).