Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Target scale shapes which gene relationships are prioritized by single-cell kinetic sensitivity

Yuchen Duan

20240601004@bucm.edu.cn

Beijing University of Chinese Medicine https://orcid.org/0009-0001-6346-6866

Research Article

Keywords: Single-cell transcriptomics, RNA velocity, kinetic Jacobian, gene regulation, perturbation

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11828610/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Target scale shapes which gene relationships are prioritized by single-cell kinetic sensitivity

Yuchen Duan

School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, China

Correspondence: 20240601004@bucm.edu.cn

Abstract

Single-cell kinetic models evaluate how a target's predicted RNA velocity responds to a small change in each candidate input while holding other modeled transcriptomic features fixed, providing a competitor baseline across all inputs to each target. We asked how target scale affects which pairs are selected and how often their targets respond in independent perturbation experiments. In a model of human CD34+ progenitors, the 50 strongest sensitivities for each of 20 candidate regulators covered 87 targets. Ranking inputs within each target increased coverage to 532 genes with column L1 scaling and 620 with percentile scaling. Three-seed sign agreement fell from 96.7% to 46.4% and 54.9%, respectively. In a genome-wide K562 CRISPRi screen, 17 of 188 evaluable raw-list pairs had a significant target response, compared with 5.60 expected under matched randomization ($q = 0.0024$). Reassigning pairs while keeping the selected targets and their frequencies fixed gave a similar count to the observed pairing ($P = 0.801$). None of the four rankings enriched for significant responses in the predicted direction across the primary K562 and CD4+ T-cell comparisons. For the leading pair by combined global ranks, ZBTB20–CENPF, standardizing late mouse thymocytes to a common cell-cycle composition reduced the estimated response from −1.07 to −0.53 log2 units. Together, these results show that target scaling governs both the breadth and sign stability of prioritized shortlists. Raw rankings preferentially recover broadly responsive targets in K562, and establishing regulator specificity and response direction requires perturbation testing in the matched cellular state.

Keywords Single-cell transcriptomics, RNA velocity, kinetic Jacobian, gene regulation, perturbation

1 Introduction

A perturbation experiment begins with a choice of regulator to perturb, target to measure and cellular setting in which to measure it. Genome-scale Perturb-seq can profile many interventions in parallel [17], but each experiment covers a defined population, perturbation modality and observation window. Perturbing the same regulator can elicit different responses as cells differentiate or encounter a stimulus. A computational screen that starts from an observed cell population can identify gene pairs worth testing and specify the direction to look for before a matched perturbation screen is available.

Single-cell kinetic models provide one way to construct such a screen [1, 2]. A kinetic Jacobian estimates how the predicted RNA velocity of a target changes after a small change in an input gene. The same calculation can be applied to thousands of gene pairs without training on measured perturbation responses. Which pairs rise to the top depends on how the sensitivities are ranked. A target with large sensitivities to many inputs can dominate several regulators' shortlists. A modest absolute sensitivity can still be unusually large relative to the other inputs to that target. These two rankings can point to different experiments.

We examined this distinction in a DynaVelo model of human hematopoietic progenitors [1, 3]. We compared absolute sensitivity with two ways of ranking each regulator relative to other inputs to the same target, and assessed the sign of each selected sensitivity across three model seeds. We used GRNBoost2 [4] to compare rankings on the discovery data and independent public perturbations to assess selected relationships in defined cellular settings. We then tested whether the rankings enriched for responsive targets, assigned them to the correct regulator and predicted their direction of response.

2 Results

2.1 Absolute sensitivity concentrates different regulators' lists on the same targets

The discovery library contained 6,035 day-7 human bone-marrow CD34+ hematopoietic stem and progenitor cells and 3,006 RNA features (GSE284047, library 8489-MV-1) [3]. We screened 1,191 candidate inputs and compared target rankings for 20 retained regulators on a common matrix evaluated in 192 cells. Selecting the 50 largest absolute sensitivities per regulator gave 1,000 edges covering only 87 targets. Thirteen targets appeared in every list. Twelve of the 13 fell in the highest joint stratum of expression mean and variance (Table S2).

For each target, we defined its scale as the sum of absolute sensitivity scores across all 1,191 inputs. Dividing each score by that sum (column L1) expanded coverage to 532 genes. Ranking inputs within each target (column percentile) expanded it to 620. Neither set of lists retained a universally shared target (Figure 1a–c).

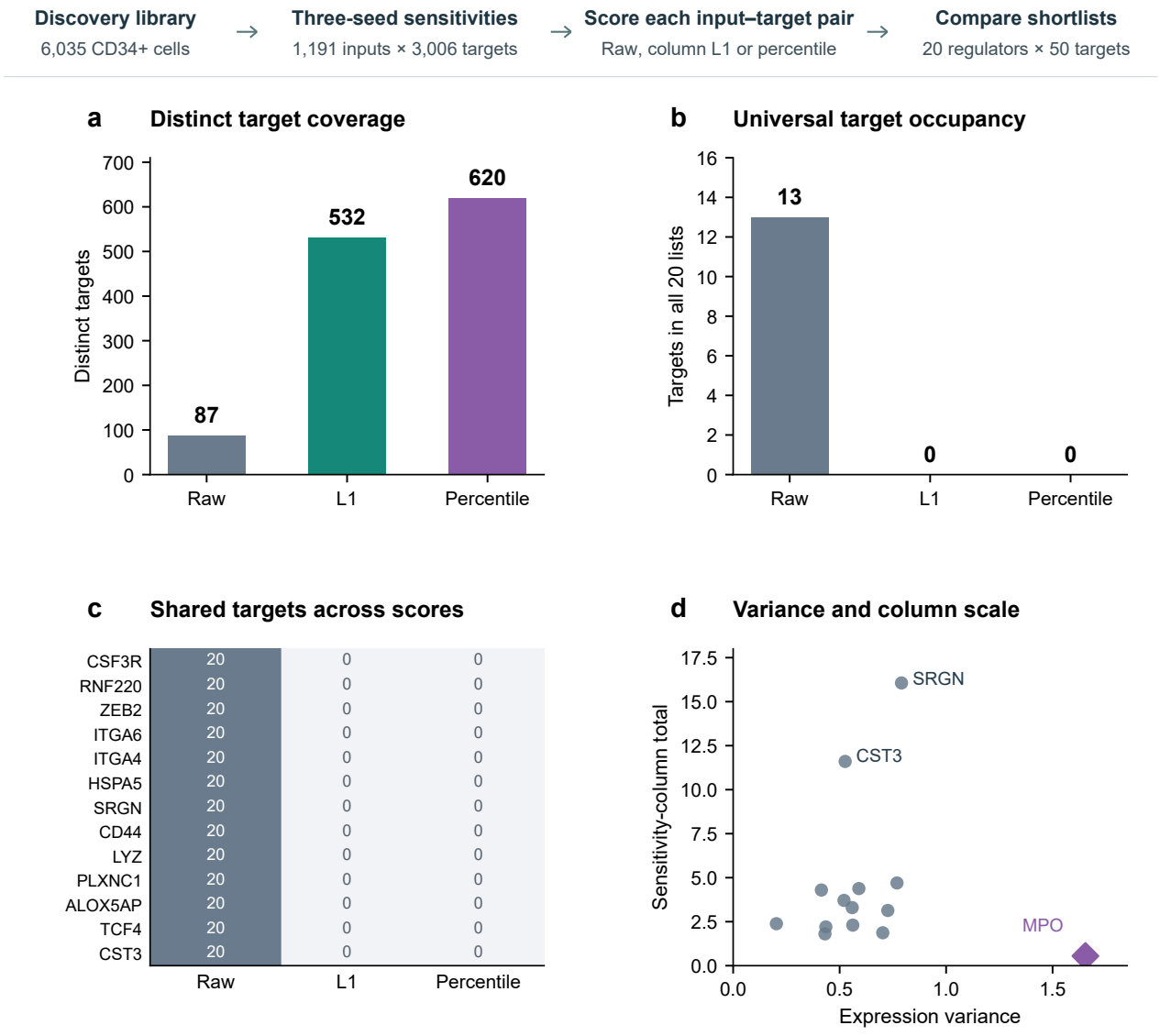

Figure 1 | Column normalization expands target coverage and resolves shared-target dominance. The workflow shows how the 20 top-50 lists were constructed (Figure S1). **a,b,** Distinct targets and targets shared by all 20 lists in the 192-cell primary matrix. **c,** Number of lists containing each of the 13 universally selected raw targets under each score. **d,** Gray circles mark these 13 targets. SRGN and CST3 have the largest sensitivity-column totals in this set. The purple diamond marks MPO. Expression variance is measured on the transformed discovery RNA scale. Each column total sums absolute sensitivity scores across 1,191 inputs (Table S2).

In the new selections, the three model seeds agreed on direction for 46.4% and 54.9% of edges, respectively, compared with 96.7% in the raw lists (Figure S2b and Table S6a). Excluding zero-containing entries prior to normalization eliminated zero selections, raised three-seed sign agreement to 65.9% and 66.4%, and increased median GRNBoost2 rank correlations to 0.4057 and 0.4062 while preserving broad target coverage (515 and 603 targets; Figure S2a,b and Table S7b). The seed values for each edge were fixed throughout. The change in agreement reflects which edges entered each shortlist.

Target scale also tracked the availability of training information. Of 3,006 targets, 598 had direct RNA-velocity supervision. Those targets accounted for 730 of the 1,000 raw selections and for 183 and 202 selections after L1 and percentile scaling, respectively (Table S8b). Across all 3,005 non-self targets per regulator, median rank correlations with GRNBoost2 [4] on the same discovery expression data remained positive under all three scores (0.4833, 0.3933 and 0.3857; Figure 2).

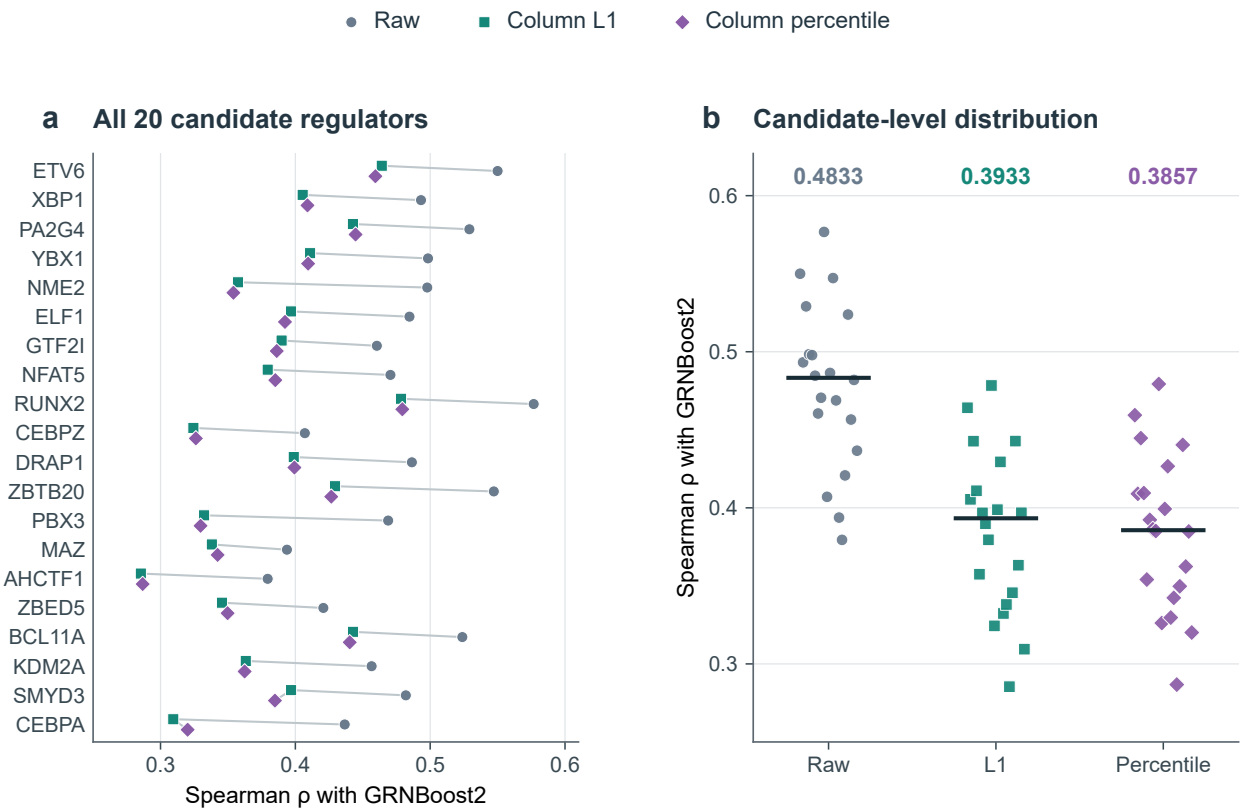

Figure 2 | Target rankings retain concordance with GRNBoost2 after column normalization. a, Each row compares three Spearman rank correlations for one candidate regulator. Segments connect scores for the same regulator, with small vertical offsets to separate the markers. Correlations use all 3,005 non-self targets, including those with zero GRNBoost2 importance, and the median raw GRNBoost2 importance across three seeds as the reference. **b,** Each point represents one regulator. Black lines and labels give the medians across the 20 regulators (Tables S1a,b).

2.2 Two CEBPA targets illustrate how absolute and relative ranks differ

CSF3R ranked 14th among CEBPA targets by absolute sensitivity, with a positive sign in all three seeds (Figure 3). Although CSF3R also appeared in every other regulator's raw top 50, it responded strongly to CEBPA induction in human K562 cells. At 48 h, CSF3R expression increased by 4.7403 log2 units relative to induced empty-vector controls (four cultures per group, adjusted P = 0.000138) [5]. A target shared across many prediction lists can therefore retain an experimentally supported relationship with an individual regulator.

MPO presents the complementary case. It had higher expression variance than any of the 13 universally selected targets, yet its sensitivity-column total was only 0.55, compared with 2.20 for CSF3R (Figure 1d and Table S2). MPO ranked 176th among CEBPA targets by absolute magnitude and joined 56 targets tied for the highest percentile score. Its signed sensitivities were −0.0022182, +0.0041965 and +0.0010788 across seeds. MPO also lacked direct RNA-velocity supervision, although its expression contributed to training. Column-percentile scaling thus surfaced MPO for follow-up despite lower directional stability than CSF3R.

In mouse progenitors, Cebpa deletion reduced Mpo expression both without stimulation (−0.5218 log2 units, adjusted P = 0.004087) and after 4 h of FLT3L stimulation (−2.1027, adjusted P = 0.009025). The same four mice per genotype contributed to both conditions [6]. The high percentile position was consistent with an established myeloid relationship [7, 8, 9]. The K562 MPO response and the mouse progenitor Csf3r responses did not reach the corresponding genome-wide adjustment threshold (Figure 4a,b).

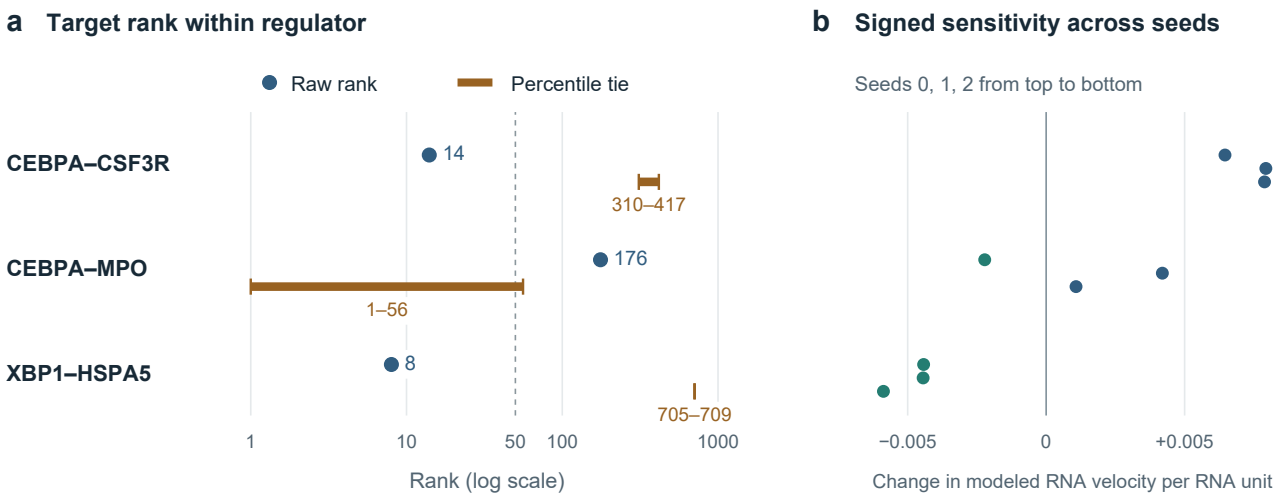

Figure 3 | A change in rank can alter the experimental shortlist while leaving the seed signs intact. a, Circles show raw ranks among targets of each named regulator. Horizontal segments show tied rank intervals after column-percentile scaling. The dashed line marks position 50. Within its highest-score tie (positions 1–56), MPO occupied position 49 by feature order and 38 with L1 tie breaking. **b,** Signed, population-weighted sensitivities from the three model seeds. Horizontal position gives the sensitivity, and small vertical offsets distinguish seeds. All ranks and signs come from the same 192-cell matrix (Tables S3d and S7d).

2.3 XBP1 loss reveals the cellular limits of directional agreement

HSPA5 illustrates why experimental context matters even for a stable model sign. It appeared in every raw list, ranked eighth for XBP1 and had a negative sensitivity in all three seeds. Conditional Xbp1 deletion increased Hspa5 in two independent mouse dendritic-cell (DC) studies, with increases of +0.3780 log2 units in splenic CD8α DCs (adjusted P = 0.000313) [10] and +0.4132 in tumor cDC1s (adjusted P = 0.008376) [11]. Both responses agree with the direction expected after loss of an input with negative sensitivity. The splenic CD11b DC estimate was close to zero (+0.0350, adjusted P = 0.7808).

In primary human CD4+ T cells stimulated for 48 h, XBP1 CRISPR interference instead reduced HSPA5 relative to non-targeting guides (−0.8179 log2 units, 95% Wald interval −1.0529 to −0.5829, adjusted P = 1.47×10^{-9}) [12]. Resting and 8-h estimates were also negative, and both had adjusted P values above 0.05 (Figure 4c). The two DC studies reproduce a net response to deletion. The human T-cell study marks a limit to transferring that direction across settings that also differ in species, stimulation and perturbation modality.

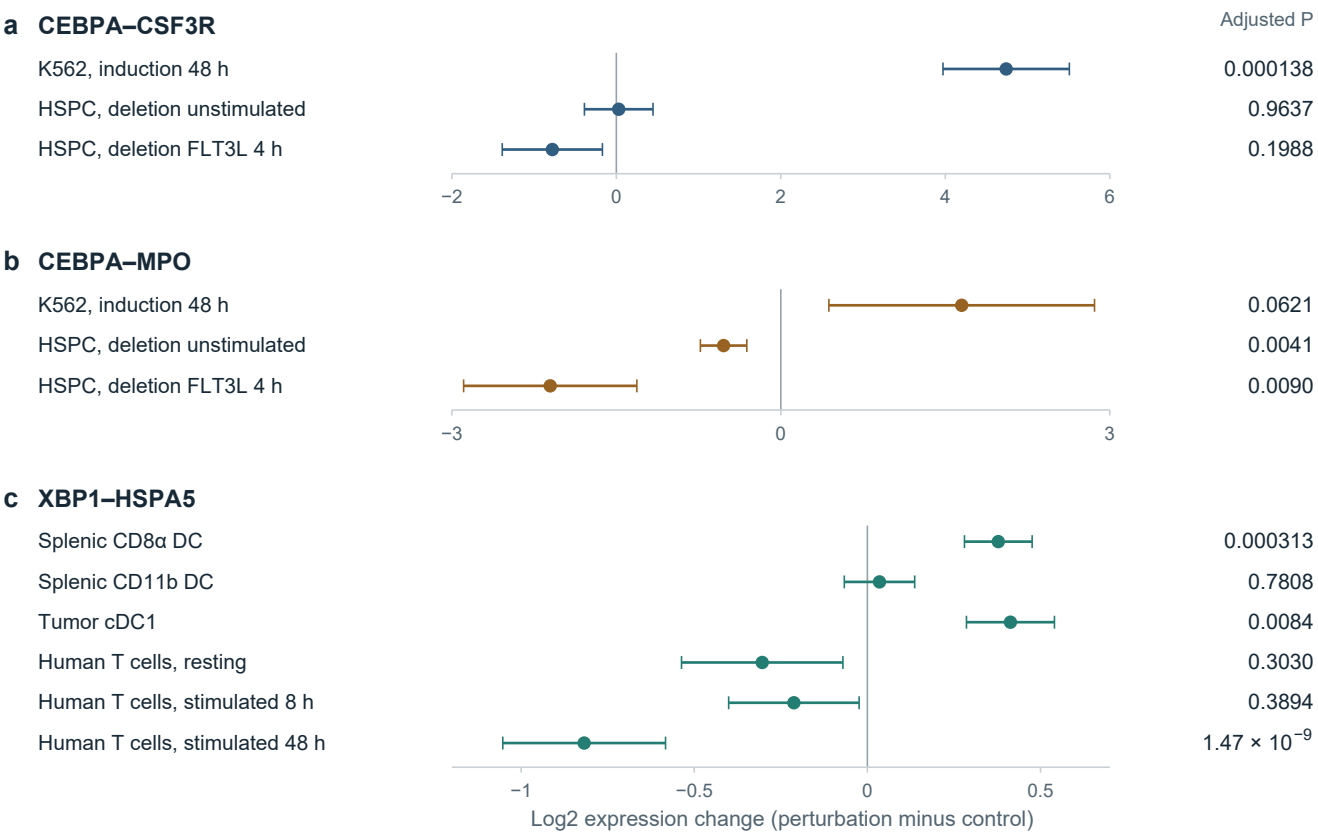

Figure 4 | Independent perturbations support different regulator–target pairs in different settings. Points show log2 expression change (perturbation minus control) with pointwise 95% intervals. The right column gives Benjamini–Hochberg adjusted P values within each contrast. **a,b,** K562 cultures (four per group) and mouse progenitors (four mice per genotype, shared across conditions). **c,** Splenic DCs (three biological replicates per genotype and subtype), tumor cDC1s (three matched wild-type and four knockout pools of two to four mice each) and primary CD4+ T cells (four-donor CRISPRi study). Panels have different horizontal scales. Table S4f identifies the published T-cell estimates and their intervals.

2.4 Joint rankings nominate gene pairs for experimental follow-up

The combined raw ranks placed ZBTB20–CENPF first among 60,100 pairs (global ranks 57 for DynaVelo and 13 for GRNBoost2). Within the 3,005 non-self ZBTB20 targets, the corresponding ranks were 2 and 8 (Table 1). Its kinetic sensitivity was positive in all three seeds. In mouse thymic organoids generated from hematopoietic stem cells [19], all three batch pairs showed lower Cenpf expression at weeks 3–4 after Zbtb20 perturbation (−0.952 log2 units, 95% CI −1.531 to −0.374, raw P = 0.00914, genome-wide BH q = 0.338). The week-2 estimate was −0.042 log2 units. Restricting the later analysis to the two batches that shared the guide and sorting protocol gave −1.287 log2 units (BH q = 0.0037). Mki67 and Top2a also declined, consistent with a possible contribution from proliferative state (Section S10). An ENCODE ZBTB20 ChIP-seq peak overlapped the CENPF transcription start site in human HepG2 cells [20].

To prioritize candidates from the relative ranking, we intersected the first 50 column L1 and GRNBoost2 target positions for each regulator. Twenty-one edges met both rank criteria. Fourteen also had the same nonzero DynaVelo sign in three seeds and were unlisted in archived CollecTRI v2.0 [18]. Table 1 shows examples from this list alongside four experimentally assessed pairs. Table S9a reports every intersection.

In genome-wide K562 CRISPRi, 17 of 188 evaluable raw-list pairs and 21 of 268 joint-list pairs showed significant target responses, exceeding the matched-randomization expectations of 5.60 and 8.51, respectively (both BH q = 0.0024). CEBPZ and DRAP1 accounted for 14 of the 17 raw hits and 15 of the 21 joint-list hits. The fixed-target tests gave P = 0.801 for the raw lists and P = 0.839 for the joint lists. The column L1 and GRNBoost2 top-50 lists shared 21 edges, compared with 17.75 expected under list permutation (P = 0.267). None of the four rankings showed significant directional enrichment in the primary K562 or CD4+ T-cell comparisons. Standardizing late thymocytes to common cell-cycle phase proportions attenuated the estimated Cenpf decrease from −1.07 to −0.53 log2 units. Among cells classified as G1-like by expression signatures, the two batches with matched guide and sorting gate gave a −1.271 log2 change (BH q = 0.0367, Section S11).

Table 1 | Prioritized candidate regulatory pairs

Regulator → target	DynaVelo rank Raw / L1 / pct.	GRNBoost2 rank	DynaVelo seed signs	External evidence or status
A. Examples assessed in independent perturbation data				
CEBPA → CSF3R	14 / 419 / 310	33	+ / + / +	Increased after CEBPA induction in human K562 cells [5].
CEBPA → MPO	176 / 190 / 49*	24	− / + / +	Decreased after Cebpa deletion in mouse progenitors under both conditions [6].
XBP1 → HSPA5	8 / 733 / 706	1	− / − / −	Increased after Xbp1 deletion in two mouse DC studies. Decreased after CRISPRi in stimulated human CD4+ T cells [10–12].
ZBTB20 → CENPF	2 / 388 / 368	8	+ / + / +	Late thymocytes, all-cell decrease (q = 0.338). G1-like decrease in batches 2–3 (q = 0.0367). HepG2 promoter occupancy [19, 20].
B. Computational candidates for follow-up				
MAZ → HMGN5	324 / 11 / 25	11	− / − / −	Unlisted in CollecTRI v2.0. K562 genome-wide, target q = 0.248.
ELF1 → WDR49	1102 / 6 / 1	28	− / − / −	Unlisted in CollecTRI v2.0. Target readout unavailable in the six screens.
DRAP1 → EVI2A	596 / 20 / 13	17	+ / + / +	Unlisted in CollecTRI v2.0. CD4+ T cells, target q range 0.474–1.000.
ELF1 → FOXO3	535 / 48 / 31	26	+ / + / +	Unlisted in CollecTRI v2.0. Confirmed ELF1 knockdown unavailable in the screens with a target readout.

Ranks are positions among 3,005 non-self targets of each regulator in the common 192-cell matrix from human CD34+ hematopoietic progenitors. GRNBoost2 ranks use three-seed median official raw importance. Signs are shown in seed order 0, 1 and 2. *MPO is in a percentile tie spanning positions 1–56. Panel B shows the first three eligible pairs by the sum of column L1 and GRNBoost2 ranks, plus ELF1→FOXO3. Table S9a shows all 21 intersections and identifies all 14 same-sign candidates.

3 Discussion

A single kinetic sensitivity can have two biological interpretations. Its magnitude describes how strongly the target responds in model coordinates. Its standing within a target column describes how prominent that input is among the inputs to the target. The CEBPA examples show why both matter. Discarding widely shared targets would remove CSF3R despite its strong response to CEBPA induction. Restricting the shortlist to the largest raw effects would miss MPO despite its decrease in both conditions of the mouse deletion experiment. MPO's position in the highest percentile tier was consistent with this experimental response. Its mixed seed signs kept it outside the sign-stable shortlist. Evaluating an input–target pair in isolation lacks this competitor baseline, making it impossible to distinguish target-wide responsiveness from regulator-specific standing.

Kinetic signs describe local responses of predicted RNA velocity to a single-input perturbation with the other modeled inputs fixed at a sampled progenitor state and can reflect covariation tied to lineage and differentiation state. Expression measured after a sustained perturbation includes feedback and changes in cell state. XBP1 can promote HSPA5 transcription [13], and Xbp1 loss in DCs can activate stress-response pathways [10, 11, 14]. The increase in Hspa5 after deletion is therefore consistent with a contribution from pathway feedback to the net response after sustained perturbation. The direction of a relationship is most informative when reported alongside the population, intervention and time at which it was measured.

Target scale, context and cell state. Our multi-screen benchmark distinguishes target responsiveness from evidence of regulator specificity. The concentration of K562 hits in CEBPZ and DRAP1 and the fixed-target permutation controls point to target composition as a major contributor to enrichment. Positive global rank concordance with GRNBoost2 coexists with a top-50 intersection near its permutation baseline. Evaluating directional accuracy in day-7 human CD34+ progenitors will require matched perturbation screens. After XBP1 loss, HSPA5 increased in mouse dendritic cells and decreased in stimulated human T cells (Section 2.3), across settings differing in lineage, stimulation and perturbation modality.

The ZBTB20–CENPF analysis illustrates how cell-state composition can shape a measured response. Phase standardization attenuated the late-thymocyte decrease, with a three-batch interval spanning zero. The G1-like reduction in the two protocol-matched batches identifies a defined state for targeted follow-up. Combining raw magnitude with column normalization can prioritize responsive targets and inputs that rank highly within those targets. Testing individual regulator–target assignments and directional hypotheses then requires perturbation in the relevant cellular state.

4 Methods

4.1 Discovery and candidate screening

DynaVelo used RNA expression, RNA-derived velocity estimates and paired ATAC-derived TF-motif accessibility from day-7 human bone-marrow CD34+ progenitors [1, 3]. Of 6,035 cells, 4,828 were used for training and 1,207 for validation. For each of 1,191 eligible inputs, a finite reduction of its RNA coordinate to its observed minimum produced a population-level perturbation score. We retained the 20 inputs with the highest median absolute-effect percentile across three seeds (Figure S1, Section S3). External perturbation responses were reserved for evaluation.

4.2 Local sensitivities and ranking

For each input–target pair, we estimated a local derivative of predicted target RNA velocity using a one-sided increment of 10^{-4} in the input's stored, log-normalized RNA coordinate. We held the other model inputs fixed. The primary matrix used 64 cells from each of three annotated strata, weighted to the proportions of the 1,884-cell evaluation population. Within each seed we averaged signed sensitivities across cells. The median absolute seed mean supplied raw magnitude. Column L1 divided this magnitude by its target-column total, and column percentile used the ascending tie-averaged rank across all 1,191 inputs. We selected the first 50 positive non-self scores per regulator, breaking ties by target feature order. Sections S1–S3 and S6–S8 give the full definitions, numerical checks and GRNBoost2 comparisons.

4.3 Candidate prioritization

We ranked 3,005 non-self targets per regulator and broke ties by frozen target feature order. The exploratory relative-score shortlist required a positive L1 score and a positive median official raw GRNBoost2 importance, both within the first 50 targets, with identical nonzero DynaVelo signs in all three seeds. We matched exact directed human gene-symbol pairs to CollecTRI v2.0 [18]. The global raw comparison ranks all 60,100 non-self edges by raw DynaVelo magnitude and by official GRNBoost2 importance, using the sum of the two ranks. See Section S9.

4.4 Perturbation evidence

We examined three relationships from the discovery rankings and the leading pair from the global raw-rank sum. We compared their model signs with expression changes under regulator induction or loss within each experimental setting. Human–mouse comparisons used reciprocally unique high-confidence orthologues. Microarray contrasts used robust, trend-adjusted limma [15]. Tumor-cDC1 counts used edgeR TMM normalization [16] and voom–limma. T-cell effects and standard errors came from the authors' CRISPRi analysis [12]. Gene-level P values were adjusted within each contrast. Sections S4–S5 give controls, experimental units, intervals and additional response comparisons. For ZBTB20–CENPF, we summed UMI counts within public thymocyte libraries [19] and fitted TMM–voom–limma models for each time point with a fixed batch term. Section S10 gives sample assignments and the ChIP-seq intersection. Section S11 describes the multi-screen benchmark, target-label permutations, and batch-paired cell-cycle sensitivity analysis using archived rankings and published expression data [12, 17, 19, 24].

Data and code availability

The MIT-licensed analysis code is available at https://github.com/dyc-max/dynavelo-target-scale-validation/releases/tag/v1.1.0. The Zenodo data archive v1.1.0, doi:10.5281/zenodo.24437365 contains model weights, sensitivity matrices and source results. The figure-source tables and scripts are in 07_source_tables_and_scripts.zip. The Section S11 reproducibility archive provides replay scripts and machine-readable results (GitHub download, Zenodo download). Study-generated weights and results are available under CC BY 4.0. Reused data retain their original terms. Supplementary Sections S1–S11 provide diagnostic tables, score definitions and source records. New public lookups and reanalyses use CollecTRI v2.0, GSE267923, and the public Perturb-seq statistics described in Section S11 [12, 17, 24].

Declarations

Funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests. The author declares no competing interests.

Ethics and consent. This study reanalyzed publicly available datasets and involved no new recruitment, human intervention or animal experiments. Original study approvals and consent procedures are described in the source publications.

Author contributions. Yuchen Duan conceived and directed the study, interpreted the results and developed the manuscript.

Use of AI-assisted tools. OpenAI Codex assisted with code development, organization of source materials, language editing and document preparation. AI tools are not authors; responsibility for the scientific content rests with the author.

References

1. Karbalayghareh A, Pelzer B, Chin CR, Melnick A, Barisic D, Leslie CS. Deep dynamical models of single-cell multiomic velocities predict loss-of-function and rescue perturbations in B cells. *Nature Biotechnology*. 2026; advance online publication. doi:10.1038/s41587-026-03348-1. Preprint: doi:10.1101/2025.04.24.650458.

2. Qiu X, Zhang Y, Martin-Rufino JD, et al. Mapping transcriptomic vector fields of single cells. *Cell*. 2022;185(4):690-711.e45. doi:10.1016/j.cell.2021.12.045.

3. Li C, Gu Y, Virgilio MC, Lee KH, Collins KL, Welch JD. Inferring differential dynamics from multi-lineage, multi-omic, and multi-sample single-cell data with MultiVeloVAE. *Nature Communications*. 2025;16:11505. doi:10.1038/s41467-025-66287-6. GSE284047.

4. Moerman T, Aibar Santos S, Bravo González-Blas C, Simm J, Moreau Y, Aerts J, Aerts S. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics*. 2019;35(12):2159-2161. doi:10.1093/bioinformatics/bty916.

5. Hughes JM, Salvatori B, Giorgi FM, Bozzoni I, Fatica A. CEBPA-regulated lncRNAs, new players in the study of acute myeloid leukemia. *Journal of Hematology & Oncology*. 2014;7:69. Article. GSE58743.

6. Anirudh S, et al. TNFα rescues dendritic cell development in hematopoietic stem and progenitor cells lacking C/EBPα. *Cells*. 2020;9(5):1223. doi:10.3390/cells9051223. GSE146288.

7. Iida S, Watanabe-Fukunaga R, Nagata S, Fukunaga R. Essential role of C/EBPα in G-CSF-induced transcriptional activation and chromatin modification of myeloid-specific genes. *Genes to Cells*. 2008;13(4):313-327. PMID:18363963.

8. Smith LT, Hohaus S, Gonzalez DA, Dziennis SE, Tenen DG. PU.1 (Spi-1) and C/EBP alpha regulate the granulocyte colony-stimulating factor receptor promoter in myeloid cells. *Blood*. 1996;88(4):1234-1247. PMID:8695841.

9. Koleva RI, et al. C/EBPα and DEK coordinately regulate myeloid differentiation. *Blood*. 2012;119(21):4878-4888. PMID:22474248.

10. Osorio F, et al. The unfolded-protein-response sensor IRE-1α regulates the function of CD8α+ dendritic cells. *Nature Immunology*. 2014;15:248-257. doi:10.1038/ni.2808. GSE52857.

11. Flores-Santibañez F, Rennen S, Fernández D, De Nolf C, Van De Velde E, Gaete González S, Fuentes C, Moreno C, Figueroa D, Lladser Á, Iwawaki T, Bono MR, Janssens S, Osorio F. Nuanced role for dendritic cell intrinsic IRE1 RNase in the regulation of antitumor adaptive immunity. *Frontiers in Immunology*. 2023;14:1209588. doi:10.3389/fimmu.2023.1209588. GSE195439.

12. Zhu R, Dann E, Yan J, Reyes Retana J, Goto R, Guitche RC, Brixi L, Ota M, Hartman A, Roth TL, Satpathy AT, Pritchard JK, Marson A. Genome-scale perturb-seq in primary human CD4+ T cells maps context-specific regulators of T cell programs and human immune traits. *Cell*. 2026;189(22):6988-7012.e17. doi:10.1016/j.cell.2026.08.002. GSE314342; author data documentation.

13. Misiewicz M, Déry MA, Foveau B, Jodoin J, Ruths D, LeBlanc AC. Identification of a novel endoplasmic reticulum stress response element regulated by XBP1. *Journal of Biological Chemistry*. 2013;288(28):20378-20391. doi:10.1074/jbc.M113.457242.

14. Tavernier SJ, Osorio F, Vandersarren L, et al. Regulated inositol-requiring enzyme 1-dependent mRNA decay sets the threshold for dendritic cell survival. *Nature Cell Biology*. 2017;19(6):698-710. doi:10.1038/ncb3518.

15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. 2015;43(7):e47. doi:10.1093/nar/gkv007.

16. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140. doi:10.1093/bioinformatics/btp616.

17. Replogle JM, Saunders RA, Pogson AN, et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. *Cell*. 2022;185(14):2559-2575.e28. doi:10.1016/j.cell.2022.05.013.

18. Müller-Dott S, Tsirvouli E, Vazquez M, et al. Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities. *Nucleic Acids Research*. 2023;51(20):10934-10949. doi:10.1093/nar/gkad841. Archived CollecTRI data, v2.0.

19. Ganapathi Sankaran D, Zhu H, Maymi VI, et al. Gene-regulatory programs that specify age-related differences during thymocyte development. *Cell Reports*. 2025;44(7):115903. doi:10.1016/j.celrep.2025.115903. GSE267923.

20. ENCODE Project. ZBTB20 ChIP-seq in human HepG2 cells. ENCSR392TWV and conservative IDR thresholded peaks, ENCFF465DSD.

21. Preiss NK, Kamal Y, Wilkins OM, et al. Characterizing control of memory CD8 T cell differentiation by BTB-ZF transcription factor Zbtb20. *Life Science Alliance*. 2023;6(9):e202201683. doi:10.26508/lsa.202201683. GSE234575.

22. Liu Q, Huang J, Zhang Z, et al. KRAS-ERK signaling drives metastasis in colorectal cancer via phosphorylation-dependent activation of the ZBTB20-TGFBR2 axis. *Journal of Experimental & Clinical Cancer Research*. 2026;45(1):29. doi:10.1186/s13046-025-03619-w. GSE309341.

23. Wang Y, Bhattacharya D. Adjuvant-specific regulation of long-term antibody responses by ZBTB20. *Journal of Experimental Medicine*. 2014;211(5):841-856. doi:10.1084/jem.20131821. GSE58659.

24. Nadig A, Replogle JM, Pogson AN, et al. Transcriptome-wide analysis of differential expression in perturbation atlases. *Nature Genetics*. 2025;57(5):1228–1237. doi:10.1038/s41588-025-02169-3. Differential-expression summary statistics, version 1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DynaVeloResearchSquareSupplementaryInformation.pdf