Supplementary Information

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

Contents. Sections S1 and S2 give the full ranking comparison and target-scale profiles. Section S3 defines the candidate screen, local sensitivity and population weights. Section S4 documents the perturbation cohorts and all target contrasts. Sections S5 and S6 report reference sensitivity and seed agreement. Sections S7 and S8 give the 17-method scoring grid and supervision audit. Supplementary Figures S1 and S2 show the study design and extended ranking diagnostics. Sections S9 and S10 give the complete candidate intersection and the ZBTB20–CENPF evidence. Section S11 reports the perturbation benchmark, overlap nulls and cell-cycle analyses. References [1–24] use the numbering of the main text.

Data and code. Model weights, full matrices and source results are available in Zenodo v1.1.0, doi:10.5281/zenodo.24437365. The analysis code is available on GitHub, release v1.1.0. The Section S11 archive provides scripts and machine-readable tables and is also available from Zenodo. Sections S1–S11 give score comparisons, perturbation estimates and records for tracing them to their sources.

Units. Sensitivity is the predicted change in RNA velocity per stored RNA-input unit. Experimental effects are log2 expression changes for perturbation minus control. Biological replication follows the original culture, animal, pool or donor design.

Figure S1. Study design and evaluation populations

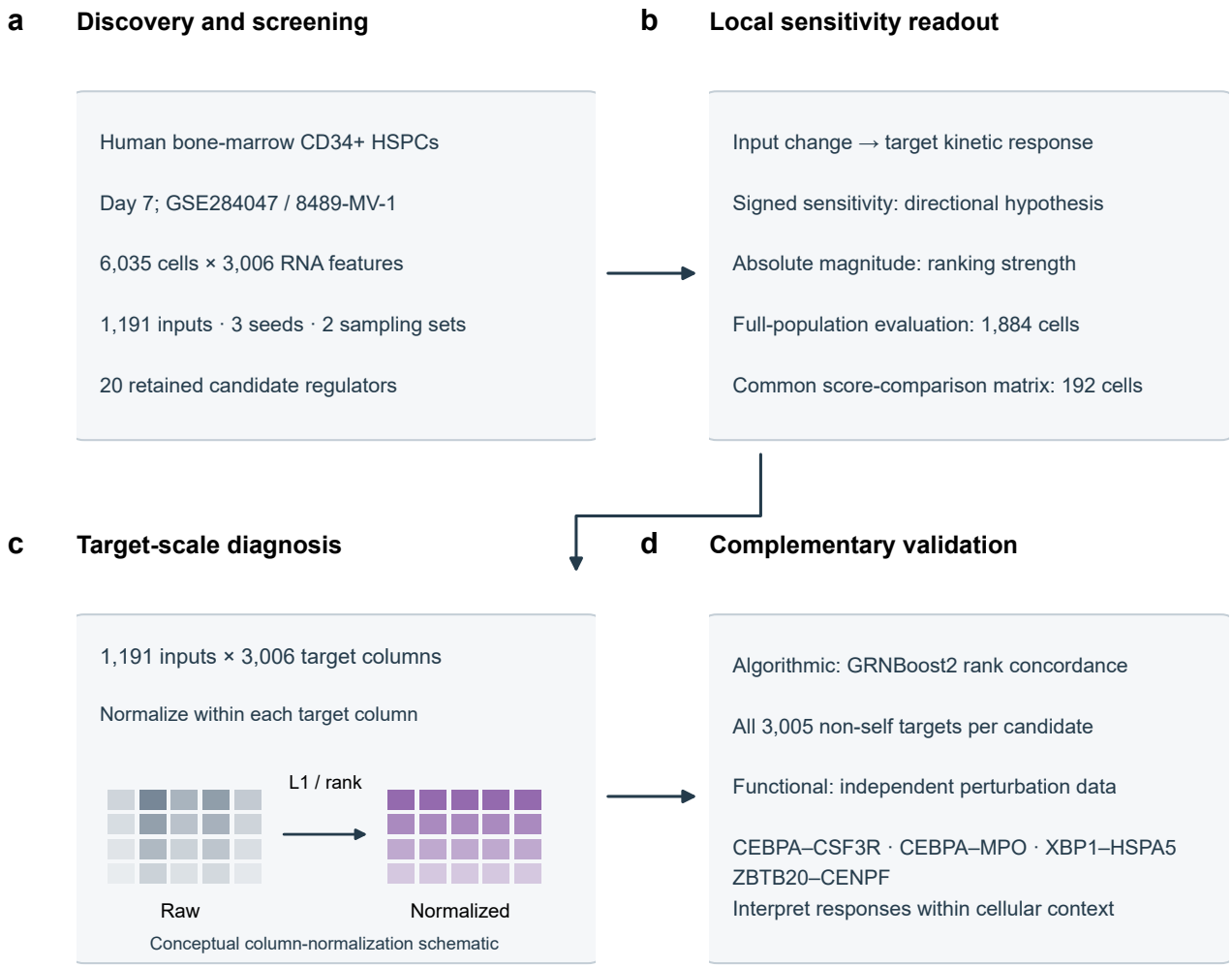


Figure S1. The discovery screen, local RNA-velocity readout and target-column transformations use a common input and target space. A balanced sample of 192 cells represents the 1,884-cell evaluation population for score comparisons. GRNBoost2 supplies an algorithmic reference on the discovery expression data. Public perturbation cohorts provide the independent response evidence. The miniature matrices illustrate the normalization operation.

S1. Target-scale comparison and all-candidate concordance

Both normalizations use the same primary matrix: 192 evaluation cells, 1,191 inputs and 3,006 targets. Each candidate contributes 50 positive-scoring non-self edges, with feature-index tie breaking. Correlations include all 3,005 non-self targets, including zero GRNBoost2 scores, and use tie-averaged ranks. The reference is the median official raw GRNBoost2 importance over three seeds. The expanded comparison against both GRNBoost2 definitions appears in S7.

Table S1a | Primary-matrix score diagnostics

Metric	Raw	Column L1	Column percentile
Distinct targets	87	532	620
Universal targets	13	0	0
Median Spearman correlation	0.483286	0.393297	0.385666
Minimum correlation	0.379467	0.285417	0.286744
Maximum correlation	0.576758	0.478379	0.479345
Reference top-50 overlaps	63	21	22

These are primary-matrix counts. The separate full-population fixed-edge comparison has 90 distinct targets, 14 universal targets and 65 official-reference overlaps (S5).

Table S1b | Full-target rank correlations for all 20 candidates

Candidate	Raw	Column L1	Column percentile
ETV6	0.549991	0.464104	0.459363
XBP1	0.493194	0.405443	0.408990
PA2G4	0.529070	0.442713	0.444650
YBX1	0.498392	0.410919	0.409460
NME2	0.497827	0.357459	0.354047
ELF1	0.484637	0.396778	0.392298
GTF2I	0.460436	0.389815	0.386189
NFAT5	0.470517	0.379504	0.385144
RUNX2	0.576758	0.478379	0.479345
CEBPZ	0.407035	0.324364	0.326195
DRAP1	0.486427	0.398823	0.399299
ZBTB20	0.547200	0.429428	0.426628
PBX3	0.468829	0.332302	0.329689
MAZ	0.393752	0.338076	0.342334
AHCTF1	0.379467	0.285417	0.286744
ZBED5	0.420763	0.345600	0.349741
BCL11A	0.523838	0.442762	0.440326
KDM2A	0.456560	0.363246	0.362379
SMYD3	0.481936	0.396824	0.384861
CEBPA	0.436597	0.309452	0.320107

Each descriptive correlation summarizes ordering across the same non-self target background.

S2. Shared-target scale profiles

Column totals sum the three-seed aggregated absolute sensitivity over all 1,191 inputs. Expression variance is the population variance of the transformed discovery-expression matrix. Joint strata are defined by expression-mean and variance quintiles. Twelve universal raw targets occupy the highest joint stratum; ALOX5AP is the exception. MPO is included for comparison and is not a universal raw target.

Table S2 | Target scale and exact top-50 occurrence

Target	Expression variance	Column total	Raw lists	L1 lists	Percentile lists
CSF3R	0.435162	2.202941	20	0	0
RNF220	0.702542	1.869695	20	0	0
ZEB2	0.561755	2.303581	20	0	0
ITGA6	0.414356	4.297879	20	0	0
ITGA4	0.769282	4.699084	20	0	0
HSPA5	0.559040	3.298778	20	0	0
SRGN	0.790581	16.061675	20	0	0
CD44	0.590291	4.380371	20	0	0
LYZ	0.726200	3.136147	20	0	0
PLXNC1	0.431088	1.806406	20	0	0
ALOX5AP	0.202995	2.384840	20	0	0
TCF4	0.520478	3.707762	20	0	0
CST3	0.525953	11.602805	20	0	0
MPO	1.654178	0.550603	0	0	1

List counts have denominator 20. The high MPO variance together with its smaller column total illustrates why variance alone does not determine raw-list inclusion. All 3,006 feature profiles are available in the deposit.

Column L1 scaling compares an input with the total sensitivity of a target to the complete candidate-input space. Percentile scaling compares its ordering within that target. The two operations therefore emphasize relative magnitude and relative position, respectively. A shared target can be biologically responsive even when it is downweighted by either operation.

S3. Extended sensitivity methods and model parameters

The discovery library was divided into 4,828 training cells and 1,207 validation cells. scVelo preprocessing used nonnegative log-normalized RNA, 30 principal components and 30 neighbors, with dynamics fitting attempted for 1,000 genes. Figure S1 traces the screening and evaluation populations.

S3.1 Input coordinate, endpoint and aggregation

Let x(c) be the stored, preprocessed RNA coordinate vector for cell c and y(c) the auxiliary input described in main Methods 4.1. For target j, input i and model seed s, the one-sided sensitivity is

$$J(c,i,j,s) = [v(j,s)(x(c) + h\, e(i), y(c)) - v(j,s)(x(c), y(c))] / h.$$

Only coordinate i is incremented. Other RNA coordinates, auxiliary inputs, model parameters and evaluation settings remain fixed; the forward model recomputes its internal quantities. The primary increment is $h = 10^{-4}$. The endpoint is predicted RNA velocity, not accumulated expression change. Units refer to the stored model coordinate and are not converted to molecules per unit time.

For finite derivative entries, the seed-specific population summary is

$$a(i,j,s) = \text{sum_c } [w(c)\, J(c,i,j,s)] / \text{sum_c } w(c).$$

The raw magnitude is $D(i,j) = \text{median_s } |a(i,j,s)|$, while the signed summary is median_s a(i,j,s). Opposite cell-level derivatives can cancel within the signed mean; absolute values are applied only at the subsequent seed aggregation step. Strict sign agreement requires every seed mean to be >0 or every seed mean to be <0. A zero, mixed signs or a nonfinite value does not satisfy this definition.

S3.2 Population weights and sampling

The full evaluation population contains 1,307 HSCs, 402 progenitor DCs and 175 DCs, totaling 1,884 cells. These broad author-defined annotations do not identify a specific external DC subtype. Primary sampling selects 64 cells without replacement from each stratum, giving 192 cells. Labels and cell identifiers are sorted before NumPy default_rng sampling; seeds are 73011 and 73012 for the primary and second sampling settings. Each sampled cell in stratum l receives weight N(l)/(1,884 x 64), with N(l) its eligible stratum size. Full-population cells have equal weight 1/1,884. Candidate selection uses a separate finite-perturbation effect, not the sum or norm of a Jacobian row. Define x_min(i) = min over all 6,035 frozen context cells of x(i). The screening operator sets x(i) to x_min(i), leaving all other RNA coordinates and motif inputs unchanged. Let delta-v(c,i,s) = v(x with x(i)=x_min(i), y) - v(x,y), and u(c,s) = v(H,c,s)/||v(H,c,s)|| the matched baseline direction on the fixed non-TF output set H. Let R(s) be the root-mean-square baseline velocity norm on H over the common baseline-evaluable full context; the scope implementation retains full-context calibration before selecting the 1,884-cell effect population. The screening endpoint is

$$E(i,s) = \text{sum_c } w(c)\, [\text{delta-v}(H,c,i,s) \text{ dot } u(c,s)] / R(s),$$

with normalized population weights. H is fixed before candidate scoring and excludes motif-associated outputs and outputs with zero training variance. Within each primary-setting seed, the average-tie ascending rank of |E| is converted to (rank - 1)/(1,191 - 1). Candidates are ordered by the median of these three percentiles, with stable gene identifier breaking ties; exactly the first 20 are retained. All 20 happen to reach percentile >=0.95 in at least two primary seeds; 19 also meet that descriptive condition in the second setting. The archived implementations are screen_scope_dispatch_v2.py and post_screen_rank_v1.py.

S3.3 Normalization and ranks

For every target column, $L(i,j) = D(i,j)/\text{sum_k } D(k,j)$, with k spanning all 1,191 inputs. The percentile score is $P(i,j) = [r(i,j) - 0.5]/1{,}191$, where r is the ascending, tie-averaged rank within the column. All-zero columns receive zero scores. Normalization includes all inputs; self-edges are excluded afterward. Exact top-50 lists contain positive scores sorted descending, with ascending original feature index breaking ties. Inclusive lists retain all scores at or above the 50th positive non-self score.

The reported full-population raw average rank is the descending, tie-averaged rank of the strict-reference three-seed aggregate magnitude. It is not an average of ranks over numerical routes. Primary-matrix CEBPA-MPO has a percentile highest-score tie of 56 targets, with deterministic index position 49 or L1-assisted tie position 38 (S7).

S3.4 Numerical scope

Full-population routes pair solver tolerance and increment as (10^{-7}, 10^{-5}), (10^{-7}, 10^{-4}), (10^{-7}, 10^{-3}) and (10^{-9}, 10^{-4}). The last is the strict reference. CEBPA-CSF3R and XBP1-HSPA5 retained their respective signs over three seeds and all four routes. The archive does not establish equivalent full numerical qualification or CPU-to-A100 equivalence for the 192-cell primary matrix. Those full-population assessments are therefore not assigned to the normalization experiment.

Score and derivative interpretation. The original column transforms include self-edges before output selection and use no positive-column floor. S7 evaluates alternative self-edge ordering, column floors, filters, Z scores and K values. Column normalization is a relative comparison in stored log-normalized coordinates. The one-sided input finite difference interrogates the complete encoder/time/dynamics/decoder forward pass; the decoder Jacobian used internally by DynaVelo is a different derivative. Existing solver- and step-size comparisons are reported in Table S3f. These describe numerical stability within the recorded routes, with model-direction stability assessed separately from external expression responses.

S3.5 Training and internal model assessment

The hidden width was 200, RNA and auxiliary latent dimensions were 50 each, batch size was 64, initial learning rate was 0.001, posterior samples numbered eight, and the maximum training duration was 200 epochs. Checkpoints were selected by each run's validation loss. The discovery library comprised 6,035 cells (4,828 training; 1,207 validation).

Input construction. The preprocessing implementation normalizes and log-transforms RNA through scVelo filter_and_normalize, computes neighborhood moments and obtains RNA velocities with recover_dynamics followed by velocity(mode='dynamical'). The recorded configuration uses 2,000 variable genes with retained eligible inputs, 30 principal components and 30 neighbors, and fits dynamics for up to 1,000 genes. Model-input assembly verifies nonnegative log-normalized RNA values and preserves unavailable velocity entries for masking; it does not apply gene-wise unit-variance scaling. ATAC peak counts are converted to chromVAR deviation z scores using hg38, JASPAR2024 human CORE motifs, motif-match P $<5 \times 10^{-5}$ and 50 GC/accessibility-matched background iterations. RNA and motif inputs are aligned by exact cell identifiers.

Optimization and integration. The archived training workflow uses Adam (with finite-value checks in the local adapter), an initial learning rate of 0.001 and ReduceLROnPlateau with patience 5, factor 0.5 and minimum learning rate 10^{-6}. Early stopping uses a patience counter of 10 and a zero improvement threshold: a strictly smaller validation loss resets the counter, a strictly larger loss increments it, and exact ties leave the counter unchanged. The model forward implementation uses the adaptive dopri5 solver; the separate trajectory-generation function uses dopri8 and should not be conflated with the assessed forward path. Checkpoints below identify the existing runs.

Internal assessment. RNA reconstruction MSE ratios compare with a training-mean baseline and are not direct measures of velocity or derivative accuracy. The reported temporal ordering is not uniformly consistent across lineages. Three model seeds measure technical stability rather than three biological replicates.

Candidate-input definition. The preprocessing audit specifies the union of Lambert HumanTFs v1.01 entries with Is TF? = Yes (Ensembl mapping) and JASPAR2024 single-TF gene assignments. Candidate genes require detection in at least 20 cells; ambiguous duplicate symbols are excluded. The RNA axis retains eligible TFs alongside 2,000 HVGs, yielding 3,006 genes and 1,191 eligible inputs. Of these inputs, 365 have selected chromVAR motifs. Candidates without a motif remain eligible for the primary RNA-only intervention. The actual audit is velocity_preparation_audit.json, plan v3; eligibility is not defined as 1,191 motif matches.

Actual velocity supervision. Although dynamics fitting was attempted for 1,000 genes, only 598 have finite RNA velocities in all 4,828 training cells; the other 2,408 have none, with no partially finite gene columns. The

RNA-velocity loss masks unavailable entries. Expression reconstruction and cross-modal velocity-consistency terms remain active for these outputs. Detailed selection counts and the HSC/progenitor-DC/DC ordering audit are in S8.

Table S3a | Training duration and selected checkpoints

Seed	Completed epochs	Selected epoch
0	90	80
1	116	106
2	137	127

Seed 0 checkpoint SHA-256: dde1af91ae8d04ec7dd918c66aacf8ee25d49843aedfc05ea967b73062e2f3c0

Seed 1 checkpoint SHA-256: 58ff1577fbe36a16033d29e30afe84638f1e501465294ae9af3a6ec40b11590c

Seed 2 checkpoint SHA-256: 70f3c44530abc5fbf09c8bef4a7d26a0f9da2bfff32668ffe29758e8007510a5

Table S3b | Internal transcriptomic and temporal assessment

Metric	Minimum across seeds	Maximum across seeds
Global RNA MSE / training-mean baseline	0.815280361	0.816583668
DC-scope RNA MSE / training-mean baseline	0.886356065	0.888698748
DC-scope temporal Spearman correlation	0.70700262	0.933701156
DC minus HSC latent time	0.219815493	0.307524681
DC minus progenitor-DC latent time	0.181550741	0.269078255
Granulocyte minus GMP latent time	-0.0788464844	-0.0550618768

Ranges describe technical model seeds. The negative granulocyte-minus-GMP time difference opposes the expected developmental order. Other archived diagnostics are retained digitally and are not used to claim validated lineage reconstruction.

Table S3c | Evaluation scopes

Scope	Cells	Purpose
Discovery library	6,035	Model training/validation and GRNBoost2 expression background
Full evaluation population	1,884	Selected relationship directions and numerical routes; fixed-edge reference comparison
Primary comparison	192	Same-matrix raw/L1/percentile scores and cross-seed sign audit

S3.6 Complete parameters for the three representative relationships

Primary-matrix signed seed means are listed independently of the full-population reference summaries. All three matrices use the same model-seed labels (0, 1 and 2).

Table S3d | Primary-matrix signed sensitivities (192 cells)

Relationship	Seed 0	Seed 1	Seed 2	Signs
CEBPA-MPO	-0.002218186301463516	0.004196475028576749	0.0010788117321639539	2 positive; 1 negative
CEBPA-CSF3R	0.00645779262359152	0.007933287133420083	0.00788571689740206	3 positive; 0 negative
XBP1-HSPA5	-0.004424440193105952	-0.00444219854919842	-0.005875306055345437	0 positive; 3 negative

CEBPA-MPO has mixed signs, unlike CEBPA-CSF3R and XBP1-HSPA5. A positive median for CEBPA-MPO does not establish three-seed direction consistency.

Table S3e(i) | Full-population reference summaries (1,884 cells)

Parameter	CEBPA-CSF3R	XBP1-HSPA5
Evaluation cells	1,884	1,884
Signed median	+0.007794984621192974	-0.004429593209064062
Raw magnitude	0.007794984621192974	0.004429593209064062
Raw average rank	15	8
Official GRNBoost2 importance	5.625221225131849	59.03482639659166
Tree-count-normalized importance	0.009583000383529555	0.10541933285105654

Only these two cases have the full-population route summaries shown in S3f. No full-population value for CEBPA-MPO is inferred.

Table S3e(ii) | Primary-matrix scores for all three cases (192 cells)

Relationship	Raw magnitude	Column L1	Column percentile
CEBPA-CSF3R	0.00788571689740206	0.003579632029138434	0.9970612930310663
XBP1-HSPA5	0.00444219854919842	0.001346619597313582	0.8249370277078085
CEBPA-MPO	0.002218186301463516	0.004028652269574075	0.9995801847187238

Values are copied from the existing candidate-target export. Signed values are in S3d. For CEBPA-MPO, membership in a 56-target highest-percentile tie is distinct from feature-index position 49 in the exact selection.

S3.7 Full-population numerical routes for two supported directions

The following signed weighted means are copied from the archived 1,884-cell relationship table. Each row denotes a technical seed and numerical route. The strict reference uses solver tolerance 10^{-9} and increment 10^{-4}.

Table S3f | Seed- and route-specific signed sensitivity

Seed	Tolerance	Increment	CEBPA-CSF3R	XBP1-HSPA5
0	10^{-7}	10^{-5}	0.006112028389204962	-0.00433260227515782
0	10^{-7}	10^{-4}	0.00611203063255055	-0.004332603149044914
0	10^{-7}	10^{-3}	0.006112053115581363	-0.004332611354254061
0	10^{-9}	10^{-4}	0.006112034788406719	-0.004332609438893687
1	10^{-7}	10^{-5}	0.007794977406834049	-0.004429583531211789
1	10^{-7}	10^{-4}	0.007794982426208025	-0.004429585324407306
1	10^{-7}	10^{-3}	0.0077950325900353985	-0.004429602836422118
1	10^{-9}	10^{-4}	0.007794984621192974	-0.004429593209064062
2	10^{-7}	10^{-5}	0.008037530147609922	-0.005863915576240122
2	10^{-7}	10^{-4}	0.008037532451406112	-0.005863915757120871
2	10^{-7}	10^{-3}	0.008037555562933268	-0.005863918199862274
2	10^{-9}	10^{-4}	0.008037537216606053	-0.00586391973813064

Definition provenance. The public release includes archived methods and a locator/checksum index for the input perturbation, population aggregation, sampling, score transformations, rank conventions, numerical routes and sign audit. Full numerical qualification remains scoped to the assessed relationship and evaluation group. The original source exports retain field names for traceability; this supplement uses their scientific definitions.

S4. Perturbation methods and designs

S4.1 Expression processing and inference

GSE58743 and GSE146288 used processed log2 expression, modeled with limma empirical Bayes estimation with robust and mean-variance trend adjustment. GSE52857 used submitted RMA-normalized expression, unique gene mappings and median log-expression summaries for multiple probes. CD8alpha and CD11b DCs were modeled separately using limma 3.68.5, with Benjamini-Hochberg adjustment over 19,901 tested genes per subtype.

GSE195439 used four Xbp1 single-knockout pools and three background-matched wild-type pools. Each pool was an independent unit comprising two to four mice. The additional wild-type pool carrying IRE1 flox was used only in sensitivity analysis; double-mutant pools were excluded. Genes required CPM >1 in at least three selected libraries. TMM normalization used edgeR 4.10.7, followed by voom and robust limma 3.68.5 with an intercept and genotype term; 10,922 genes were tested. Pool-specific mouse identities and batch assignments were not available in the retained metadata.

Human targets were linked to mouse genes through high-confidence, reciprocally unique one-to-one orthologues. Effects always represent perturbation minus control. Genome-wide Benjamini-Hochberg values refer to the tested genes within each contrast. For the original CEBPA target family, Holm correction retained all six contrasts and two targets (12 endpoints). Targets lacking replicate-level variance remained in that family with conservative correction inputs; no variance or inferential P value was inferred from group means. Ninety available individual expression measurements are consolidated in Table S4b. THP1 contrasts have group-level summaries only.

S4.2 Replication and contrasting contexts

K562 biological units were reported independent cultures, with no pairing between constructs. The two HSPC treatment conditions used the same eight mice and were compared separately. Splenic DC units were the three reported biological replicates per genotype within each subtype; matching replicate labels across subtypes was not treated as proof of independent additional animals. T-cell statistics follow the source authors' CRISPRi analysis of targeting-guide versus non-targeting-guide pseudobulks separately in resting, 8-h and 48-h stimulation conditions, with donor and aggregated-cell-number covariates [12]. The study contained four donors; overlapping two-donor summaries are not independent additional cohorts.

Multiplicity formulas. For a family of m unadjusted values sorted as p(1) <= ... <= p(m), BH-adjusted values are q(i) = min[1, min over k >= i of m p(k)/k]. Holm-adjusted values are h(i) = min[1, max over k <= i of (m - k + 1) p(k)]. Values are returned to their original order. Gene-level BH families contain all tested genes within a contrast, rather than only the displayed targets. The CEBPA Holm family retains all 12 target endpoints. Available model-based confidence intervals are marginal 95% intervals, not simultaneous or BH-adjusted intervals; excluding zero does not necessarily imply BH significance.

Interval recovery and verification. The two GSE52857 models were rerun only on the already prepared 19,901-gene expression matrix with the identical intercept-plus-genotype design and robust/trend empirical Bayes settings. The maximum absolute discrepancy in logFC, moderated t, P and BH values was required to be $<10^{-8}$ against every archived gene result before topTable(confint=TRUE) was used. SE(j) = stdev.unscaled(j) sqrt(s2.post(j)); the interval is effect(j) ± t(0.975, df.total(j)) SE(j). Neither probe processing nor permutation analyses were repeated. T-cell intervals are author log_fc ± 1.959964 lfcSE, using the published DESeq2 estimates rather than a local refit.

Table S4a | Cohort and contrast overview

Accession / context	Control / altered units	Perturbation and timing
GSE58743; human K562	4 / 4 cultures	CEBPA induction vs induced empty vector; 48 h
GSE146288; mouse HSPC	4 / 4 mice, shared across conditions	Cebpa deletion vs WT; unstimulated or FLT3L 4 h
GSE52857; mouse splenic DC	3 / 3 reported biological replicates per subtype	Conditional Xbp1 deletion vs WT; subtypes separate
GSE195439; mouse tumor cDC1	3 / 4 independent pools	Xbp1 single deletion vs background-matched WT
GSE314342; human CD4 T cells	Four donors in source study	XBP1 CRISPRi vs non-targeting guides; resting, 8-h and 48-h stimulation
GSE84388; mouse PreGM	3 / 3 individual mice	Cebpa deletion contrast; 18 days after pIC; manipulation readout unresolved
GSE119551; human THP1	3 / 3 reported per contrast	CEBPA knockdown; solvent or vitamin D; individual variance unavailable

S4.3 CEBPA sample assignments and individual expression

GSE58743: K562 induction at 48 h. GSE84388: PreGM, 18 days after pIC. GSE146288: HSPCs, unstimulated or after FLT3L for 4 h. Mouse identifiers repeat across HSPC conditions; no between-genotype pairing is assumed. Batch assignments were unreported. Expression entries are the archived processed log2 values, arranged by sample without aggregation.

Table S4b | Sample-level CEBPA, MPO and CSF3R measurements

Sample	Context	Group	Unit ID	CEBPA	MPO	CSF3R
GSM1418992	K562, 48 h	Control	Rep. 1	6.638000	1.705000	7.911000
GSM1418993	K562, 48 h	Control	Rep. 2	8.229000	1.701000	8.653000
GSM1418994	K562, 48 h	Control	Rep. 3	7.784000	1.692000	9.376000
GSM1418995	K562, 48 h	Control	Rep. 4	8.440000	1.664000	9.201000
GSM1418996	K562, 48 h	Induced	Rep. 1	9.090000	1.689000	13.885000
GSM1418997	K562, 48 h	Induced	Rep. 2	9.286000	3.994000	13.228000
GSM1418998	K562, 48 h	Induced	Rep. 3	9.748000	3.703000	13.502000
GSM1418999	K562, 48 h	Induced	Rep. 4	9.712000	3.974000	13.487000
GSM2232912	PreGM, day 18	Control	Rep. 1	6.934015	12.428972	6.948210
GSM2232913	PreGM, day 18	Control	Rep. 2	6.711763	12.359520	6.705489
GSM2232914	PreGM, day 18	Control	Rep. 3	6.845897	12.436413	6.934791
GSM2232915	PreGM, day 18	KO	Rep. 1	8.708371	12.852708	7.598983
GSM2232916	PreGM, day 18	KO	Rep. 2	9.310577	12.960272	8.107243
GSM2232917	PreGM, day 18	KO	Rep. 3	8.983926	12.848520	7.759067
GSM4376759	HSPC, FLT3L 4 h	KO	515	5.133750	9.046540	5.978380
GSM4376760	HSPC, FLT3L 4 h	KO	581	4.890360	7.351460	5.127820
GSM4376761	HSPC, FLT3L 4 h	KO	584	4.449860	7.904720	5.921720
GSM4376762	HSPC, FLT3L 4 h	KO	608	4.677620	9.076140	5.815010
GSM4376763	HSPC, unstimulated	KO	515	4.627340	11.239400	7.490700
GSM4376764	HSPC, unstimulated	KO	581	5.387470	10.954800	6.767460
GSM4376765	HSPC, unstimulated	KO	584	4.816530	11.093700	7.324270
GSM4376766	HSPC, unstimulated	KO	608	5.038980	11.021300	7.450010
GSM4376767	HSPC, FLT3L 4 h	Control	536	6.893800	10.360300	5.697430
GSM4376768	HSPC, FLT3L 4 h	Control	567	6.698770	10.663600	6.646220
GSM4376769	HSPC, FLT3L 4 h	Control	570	7.668440	10.629700	6.931750
GSM4376770	HSPC, FLT3L 4 h	Control	606	7.051570	10.136200	6.680240
GSM4376771	HSPC, unstimulated	Control	536	7.439500	11.752000	7.517490
GSM4376772	HSPC, unstimulated	Control	567	7.535800	11.698100	6.984880
GSM4376773	HSPC, unstimulated	Control	570	7.665410	11.663300	7.435010
GSM4376774	HSPC, unstimulated	Control	606	7.248250	11.283100	6.978970

Uppercase symbols label the cross-species relationship; mouse entries correspond to Cebpa, Mpo and Csf3r. Sample accession identifies the individual PreGM unit. No individual THP1 values were available.

S4.4 Complete CEBPA contrast results

Six contrasts and all three measured genes are retained below. C1: K562 induction (GSE58743). C2: PreGM deletion (GSE84388). C3: HSPC deletion, unstimulated (GSE146288). C4: HSPC deletion, FLT3L 4 h (GSE146288). C5: THP1 knockdown, solvent (GSE119551). C6: THP1 knockdown, vitamin D (GSE119551). The CEBPA self-readout is a manipulation check and is outside the 12-target-endpoint Holm family.

Table S4c | Effects, uncertainty and multiple-testing values

Contrast	Gene	log2 change	95% model interval	P	BH-adjusted P	Holm-adjusted P
C1	CEBPA	1.6863	0.7517 to 2.6208	0.003538	0.022245	Not in family
C1	MPO	1.6495	0.4383 to 2.8607	0.014380	0.062067	0.100663
C1	CSF3R	4.7403	3.9719 to 5.5086	1.237e-06	0.000138	1.485e-05
C2	Cebpa	2.1704	1.8621 to 2.4787	5.538e-09	5.510e-06	Not in family
C2	Mpo	0.4789	0.3646 to 0.5931	1.329e-06	0.000264	1.485e-05
C2	Csf3r	0.9589	0.6704 to 1.2474	1.301e-05	0.001233	0.000130
C3	Cebpa	-2.5047	-2.8950 to -2.1143	5.948e-09	2.753e-06	Not in family
C3	Mpo	-0.5218	-0.7336 to -0.3101	0.000164	0.004087	0.001311
C3	Csf3r	0.0290	-0.3881 to 0.4462	0.882393	0.963689	1.000000
C4	Cebpa	-2.2902	-2.7903 to -1.7902	4.328e-07	0.000187	Not in family
C4	Mpo	-2.1027	-2.8928 to -1.3127	0.000134	0.009025	0.001204
C4	Csf3r	-0.7782	-1.3871 to -0.1693	0.016610	0.198776	0.100663
C5	CEBPA	-0.8659	Unavailable	Unavailable	Unavailable	Not in family
C5	MPO	0.2031	Unavailable	Unavailable	Unavailable	Not inferential
C5	CSF3R	0.1499	Unavailable	Unavailable	Unavailable	Not inferential
C6	CEBPA	-1.0837	Unavailable	Unavailable	Unavailable	Not in family
C6	MPO	-2.3876	Unavailable	Unavailable	Unavailable	Not inferential
C6	CSF3R	-0.5805	Unavailable	Unavailable	Unavailable	Not inferential

For THP1, retained changes derive from exported group means; original units and replicate variance were unresolved. The adjustment procedure used conservative P = 1 placeholders for those target endpoints, but these are not measured inferential P values.

In PreGM cells, target directions differ from the positive myeloid examples, while the Cebpa assay readout increases in the deletion contrast. The manipulation readout is therefore unresolved for interpreting that context. In K562, the positive MPO estimate has BH-adjusted P = 0.062067 and Holm-adjusted P = 0.100663. HSPC Csf3r is near zero without stimulation and negative after FLT3L, but neither contrast passes the retained multiplicity corrections. These results distinguish context limitations and incomplete evidence from a uniform absence of response.

S4.5 DC sample assignments

Splenic DC accessions belong to GSE52857; tumor cDC1 accessions belong to GSE195439. WT denotes wild type, KO denotes Xbp1 single knockout and double mutant denotes the combined Xbp1/IRE1 perturbation. Tumor pools are independent biological units; animal IDs within pools and batch assignments are unavailable.

Table S4d | Splenic DC biological-replicate mapping

Sample	Subtype	Genotype	Replicate label
GSM1277090	CD8alpha DC	WT	1
GSM1277091	CD8alpha DC	WT	2
GSM1277092	CD8alpha DC	WT	3
GSM1277093	CD8alpha DC	KO	1
GSM1277094	CD8alpha DC	KO	2
GSM1277095	CD8alpha DC	KO	3
GSM1277096	CD11b DC	WT	1
GSM1277097	CD11b DC	WT	2
GSM1277098	CD11b DC	WT	3
GSM1277099	CD11b DC	KO	1
GSM1277100	CD11b DC	KO	2
GSM1277101	CD11b DC	KO	3

Table S4e | Tumor cDC1 inclusion and control background

Sample	Genotype	Group	Primary	Sensitivity
GSM5836786	XBP1 fl/fl	Matched WT	Yes	Yes
GSM5836787	XBP1 fl/fl	Matched WT	Yes	Yes
GSM5836788	XBP1 fl/fl	Matched WT	Yes	Yes
GSM5836789	XBP1 fl/fl, IRE1 fl/fl	Extra WT; IRE1 flox	No	Yes
GSM5836790	XBP1 fl/fl, CD11c-CRE	Xbp1 single KO	Yes	Yes
GSM5836791	XBP1 fl/fl, CD11c-CRE	Xbp1 single KO	Yes	Yes
GSM5836792	XBP1 fl/fl, CD11c-CRE	Xbp1 single KO	Yes	Yes
GSM5836793	XBP1 fl/fl, CD11c-CRE	Xbp1 single KO	Yes	Yes
GSM5836794	XBP1 fl/fl, IRE1 fl/fl, CD11c-CRE	Double mutant	No	No
GSM5836795	XBP1 fl/fl, IRE1 fl/fl, CD11c-CRE	Double mutant	No	No
GSM5836796	XBP1 fl/fl, IRE1 fl/fl, CD11c-CRE	Double mutant	No	No

The primary contrast uses only matched WT and Xbp1 single-KO pools (3 vs 4). The extra WT enters a separate sensitivity comparison. Double mutants enter neither comparison.

S4.6 XBP1-HSPA5 effects across all retained contexts

DC effects come from the local reanalyses specified above. T-cell values were directly verified against the authors' GWCD4i.DE_stats.h5ad artifact [12], with all three culture conditions retained. Their DESeq2 pseudobulk model includes donor, log10 aggregated cell count and perturbation, with non-targeting guides as controls. The four-donor study is the biological source; cells and overlapping donor subsets are not additional independent experiments. All intervals below are pointwise and can exclude zero when the across-gene BH value exceeds 0.05.

Table S4f | HSPA5 response across contexts

Context	log2 change	95% interval	Raw P	BH q
CD8alpha DC	0.3780	0.2803 to 0.4758	2.53032e-06	0.000312769
CD11b DC	0.0350	-0.0666 to 0.1366	0.461488	0.780825
Tumor cDC1, 3 WT / 4 KO	0.4132	0.2861 to 0.5402	2.31075e-05	0.00837574
Tumor cDC1, 4 WT / 4 KO	0.3985	0.2902 to 0.5069	4.68658e-06	0.00183731
T cells, resting	-0.3037	-0.5368 to -0.0707	0.0106419	0.302979
T cells, stimulated 8 h	-0.2122	-0.4006 to -0.0238	0.0272481	0.389373
T cells, stimulated 48 h	-0.8179	-1.0529 to -0.5829	9.0598e-12	1.4689e-09

The 4-WT tumor comparison adds one IRE1-flox control and is a sensitivity analysis of the same experiment, not an independent replication. The primary 3-WT result is shown in Figure 4c.

The tumor-cDC1 Hspa5 standard error in the primary comparison is 0.0573392728912815. GSE52857 Hspa5 standard errors are 0.0447434708265477 (CD8alpha DC) and 0.0457283811718172 (CD11b DC), with moderated degrees of freedom 11.712065 and 10.181524. For K562 CSF3R, untreated HSPC Mpo and FLT3L-treated HSPC Mpo, the original Holm-adjusted values across 12 target endpoints are 0.0000148492, 0.00131096 and 0.00120407.

T-cell source identification. Artifact: https://genome-scale-tcell-perturb-seq.s3.amazonaws.com/marson2025_data/GWCD4i.DE_stats.h5ad. Retrieved 17 November 2026 by bounded byte-range reads. Object ETag: c9ff52fcc6d6ce8a387a76dc757a5b97-2002; object size: 16,786,240,107 bytes. XBP1 rows are 31627 (Rest), 31628 (Stim48hr) and 31629 (Stim8hr), and HSPA5 is column 401 (zero-based). Source layers are log_fc, p_value, adj_p_value and lfcSE. The three target symbols and conditions were verified from the object's metadata. Table S4f reports the expression estimates recovered from these rows. The ETag identifies the source object. Author documentation: https://github.com/emdann/GWT_perturbseq_analysis_2025/blob/master/metadata/data_sharing_readme.md.

Contextual interpretation. XBP1 overexpression and promoter binding support positive HSPA5 transcriptional regulation [13]. DC studies report IRE1 hyperactivation and ER-stress responses after Xbp1 loss [10, 11, 14]. These findings motivate feedback as an interpretation of the DC net increase, while an ATF6-specific cause for the analyzed contrasts remains unassigned. No direct repressive mechanism is inferred from the negative model derivative.

Additional datasets. GSE84388 and GSE119551 supply the PreGM and THP1 comparisons in Tables S4a-c. PreGM interpretation depends on the unresolved manipulation readout. THP1 group summaries lack replicate-level variance. These are data-interpretation limits, distinct from a statistically nonsignificant replicated effect.

S5. Reference sensitivity and cross-study response evidence

S5.1 Fixed-edge reference comparison

This assessment uses the 1,884-cell full-population matrix, not the 192-cell primary matrix. Its original 1,000 edges cover 90 distinct targets, 14 universally shared targets and 65 overlaps with official raw GRNBoost2 top-50 lists. The numerically qualified subset contains 952 edges. Both edge sets remain fixed when the reference score is changed.

Target-column permutations preserve shared target membership across candidate lists and operate within 25 joint expression-mean/variance quintile strata. The 20 candidate self-target columns are held fixed; 2,986 columns are permuted. The archived calculation used 99,999 permutations with random seed 20261116. P values include the observed configuration through a plus-one correction. Candidate-specific BH adjustment applies within each 20-candidate comparison family. These enrichment tests ask a different question from full-target rank correlation.

Candidate-level and pooled nulls. The candidate-level values in S5b-c use a 25-stratum hypergeometric convolution: within each stratum h, X(h) follows Hypergeometric(N(h), K(h), n(h)), where N is the eligible target background, K is reference top-50 membership and n is the candidate's selected-target count. The one-sided enrichment P is Pr[sum_h X(h) >= observed overlap]. The pooled result instead uses common target-column permutations across candidates to preserve shared membership, with P = (1 + number of randomized overlaps at least as large as observed)/(99,999 + 1). These are different tests, and candidate-level P values are not the pooled permutation P. BH adjustment is applied separately within each 20-candidate family.

Qualification scope. The original 1,000 and qualified 952 sets are both retained. The 48-edge attrition is not a biological rejection count. This readable source collection does not supply a complete per-edge reason decomposition, so the manuscript does not assign an unverified failure cause or transfer qualification to the separate primary matrix.

Table S5a | Dependence-preserving pooled reference sensitivity

Fixed edge set	Reference importance	Overlap	Permutation P
Original 1,000	Official raw	65	0.00121
Original 1,000	Tree-count normalized	25	0.40206
Qualified 952	Official raw	60	0.00314
Qualified 952	Tree-count normalized	23	0.46994

Official raw and tree-count-normalized importance use different GRNBoost2 score definitions. Dividing by fitted tree count changes ordering across targets. The fixed-edge comparison measures sensitivity to that reference choice; S7 additionally compares every primary-matrix scoring method against both definitions on an identical target background.

S5.2 Original 1,000 edges

Table S5b | Candidate-level reference comparison

Candidate	Edges	Raw overlap	Raw P	Raw BH	Normalized overlap	Normalized P	Normalized BH
ETV6	50	6	0.050688	0.268566	4	0.183652	1.000000
XBP1	50	4	0.126011	0.330320	2	0.310586	1.000000
PA2G4	50	1	0.944375	0.994079	1	0.845212	1.000000
YBX1	50	3	0.360380	0.600633	0	1.000000	1.000000
NME2	50	1	0.921601	0.994079	1	0.798255	1.000000
ELF1	50	1	0.881970	0.994079	1	0.523303	1.000000
GTF2I	50	2	0.506498	0.675331	0	1.000000	1.000000
NFAT5	50	4	0.080570	0.268566	0	1.000000	1.000000
RUNX2	50	5	0.148644	0.330320	1	0.853753	1.000000
CEBPZ	50	3	0.214926	0.429852	1	0.449822	1.000000
DRAP1	50	3	0.392530	0.603892	0	1.000000	1.000000
ZBTB20	50	6	0.068385	0.268566	2	0.734580	1.000000
PBX3	50	0	1.000000	1.000000	0	1.000000	1.000000
MAZ	50	5	0.001944	0.038888	3	0.004043	0.080869
AHCTF1	50	6	0.005647	0.056471	1	0.550619	1.000000
ZBED5	50	2	0.493648	0.675331	1	0.241957	1.000000
BCL11A	50	4	0.311677	0.566685	1	0.813097	1.000000
KDM2A	50	3	0.141682	0.330320	1	0.569437	1.000000
SMYD3	50	1	0.940567	0.994079	1	0.872817	1.000000
CEBPA	50	5	0.062881	0.268566	4	0.047891	0.478913

Raw and normalized refer to GRNBoost2 importance. P values belong to candidate target sets; they are not edge-level tests. BH adjustment is within each 20-candidate family.

S5.3 Numerically qualified 952 edges

Table S5c | Candidate-level reference comparison

Candidate	Edges	Raw overlap	Raw P	Raw BH	Normalized overlap	Normalized P	Normalized BH
ETV6	49	5	0.125210	0.314849	3	0.382178	1.000000
XBP1	48	4	0.111064	0.314849	2	0.291672	1.000000
PA2G4	47	1	0.935781	0.990070	1	0.831055	1.000000
YBX1	49	3	0.345183	0.627605	0	1.000000	1.000000
NME2	47	1	0.909962	0.990070	1	0.780852	1.000000
ELF1	44	1	0.840715	0.990070	1	0.474083	1.000000
GTF2I	49	2	0.493843	0.658458	0	1.000000	1.000000
NFAT5	48	4	0.070222	0.280888	0	1.000000	1.000000
RUNX2	48	5	0.129091	0.314849	1	0.840349	1.000000
CEBPZ	47	2	0.466303	0.658458	1	0.436083	1.000000
DRAP1	50	3	0.392530	0.654217	0	1.000000	1.000000
ZBTB20	45	6	0.043805	0.280888	2	0.678352	1.000000
PBX3	46	0	1.000000	1.000000	0	1.000000	1.000000
MAZ	48	4	0.011985	0.239698	2	0.046085	0.460852
AHCTF1	46	4	0.069592	0.280888	1	0.520195	1.000000
ZBED5	46	2	0.456406	0.658458	1	0.232599	1.000000
BCL11A	48	4	0.280460	0.560921	1	0.798585	1.000000
KDM2A	50	3	0.141682	0.314849	1	0.569437	1.000000
SMYD3	50	1	0.940567	0.990070	1	0.872817	1.000000
CEBPA	47	5	0.050601	0.280888	4	0.039711	0.460852

Raw and normalized refer to GRNBoost2 importance. P values belong to candidate target sets; they are not edge-level tests. BH adjustment is within each 20-candidate family.

S5.4 DC cross-study concordance

The common-target comparison links GSE195439 tumor cDC1 responses to each GSE52857 splenic subtype. Both studies must have scored the target. Effects are knockout minus wild type. The following correlations and sign counts compare experiment with experiment, not kinetic signs with perturbation direction. Only one CD8alpha target and no CD11b target met gene-level adjusted P <0.05 in both studies.

Table S5d | Cross-study response summary

Splenic subtype	Common targets	Same-sign responses	Spearman correlation	Adjusted P <0.05 in both
CD8alpha DC	27	19/27	0.4939	1
CD11b DC	27	20/27	0.5366	0

Table S5e | All 27 commonly scored targets

Human orthologue	Tumor log2FC	Tumor BH	CD8alpha log2FC	CD8alpha BH	CD11b log2FC	CD11b BH
ITGA4	0.1292	0.937461	-0.1169	0.164158	0.0006	0.996931
CD44	-0.2252	0.797913	-0.2272	0.056537	-0.0212	0.923784
ITGA6	-0.0912	0.956028	-0.0095	0.992403	-0.1416	0.462030
TCF4	-0.2617	0.805566	-0.1112	0.893771	-0.0751	0.675711
HSPA5	0.4132	0.008376	0.3780	0.000313	0.0350	0.780825
ZEB2	0.1967	0.982599	-0.5154	0.025407	-0.0537	0.704516
ALOX5AP	-0.1082	0.826257	0.0163	0.910540	-0.0555	0.606465
TOP2A	0.1368	0.963290	0.1981	0.344983	0.0951	0.665756
STOM	-0.0927	0.962094	-0.3135	0.003997	-0.1095	0.464692
GLIPR1	-0.1551	0.845741	0.0127	0.955465	-0.0018	0.991656
RNF220	0.0275	0.991050	0.1048	0.371274	-0.0931	0.540777
ITM2A	0.1414	0.989973	-0.9108	0.346405	-0.3321	0.427545
PLXNC1	0.2468	0.891812	-0.2998	0.193847	-0.0057	0.975338
PLXDC2	0.2244	0.976133	0.0242	0.946997	0.1064	0.578892
SORL1	-0.5196	0.536347	-0.7886	0.004441	-0.0944	0.504466
HDAC9	0.0655	0.981453	0.3305	0.010330	-0.0051	0.987726
ANGPT2	5.0929	0.805566	0.0936	0.666929	0.0484	0.848098
CTR9	0.1131	0.963797	-0.0523	0.833665	-0.1198	0.510837
IRF8	-0.0175	0.992470	0.0334	0.769848	-0.0841	0.876091
MED12L	0.7034	0.631574	0.2287	0.180192	0.1717	0.464692
WDFY2	-0.0652	0.991778	-0.1821	0.117413	-0.0695	0.719870
PRSS57	-0.9039	0.159747	-0.4511	0.008452	-0.0891	0.658860
ALCAM	-0.2293	0.549099	-0.0025	0.993137	-0.0408	0.816898
IL6R	-0.2753	0.523033	-0.4259	0.014005	-0.0621	0.752773
PRDX1	-0.0547	0.963290	-0.0748	0.458615	0.0254	0.867926
SELL	-0.7756	0.327891	-1.2046	0.150021	-0.4118	0.484237
RUNX1	-0.1566	0.797913	-0.0833	0.523155	-0.0647	0.609056

Mouse expression effects are labeled by the mapped human symbol. Values are copied from the two archived subtype comparisons. The digital source also retains mapping and expression-filter coverage for targets outside the common scored set.

S5.5 Primary tumor-cDC1 target-set ranking

Candidate target-set scores use background percentiles of absolute moderated t statistics, centered within control-expression mean-variance strata and summarized across scorable targets. The primary target family is restricted to numerically qualified relationships. The primary permutation statistic is the XBP1 score minus the mean score of the other 19 candidate sets, evaluated over all 35 genotype-label allocations of the seven independent pools. This is a response-strength ranking, not a model-direction agreement rate.

Table S5f | Primary unremoved comparison for all 20 candidates

Candidate	Scored targets	Centered score	Uncentered score	Rank	Genes BH <0.05
ETV6	28	0.070865	0.574656	19	1
XBP1	27	0.123773	0.631388	1	1
PA2G4	27	0.101530	0.606163	5	1
YBX1	27	0.117946	0.633231	2	1
NME2	27	0.071457	0.578343	17	1
ELF1	26	0.095298	0.605838	6	1
GTF2I	27	0.077071	0.583752	13	1
NFAT5	28	0.088804	0.592595	7	1
RUNX2	26	0.087293	0.590046	9	1
CEBPZ	26	0.082685	0.583850	11	1
DRAP1	26	0.083801	0.587200	10	1
ZBTB20	26	0.070904	0.575084	18	1
PBX3	27	0.115623	0.626125	3	1
MAZ	28	0.073817	0.577159	14	1
AHCTF1	26	0.088365	0.589064	8	1
ZBED5	25	0.082271	0.582489	12	1
BCL11A	26	0.102723	0.613218	4	1
KDM2A	31	0.073298	0.569819	15	1
SMYD3	31	0.071927	0.564664	16	2
CEBPA	27	0.065112	0.564833	20	1

XBP1 ranks 1/20; its primary permutation P is 0.0857142857142857 (35 label allocations). The best rank is not equivalent to statistical significance.

An archived summary reported rank 1/20 after removing HSPA5 (26 remaining targets) and rank 1/20 after removing eight shared targets (19 remaining), with a splenic-DC shared-target-removal rank of 6/20. The retained collection lacks the eight-target identities and complete deletion-specific candidate scores, denominators and permutation results. These narrative summaries are preserved for traceability but excluded from the main evidential argument. The primary P = 0.0857 applies only to the original 27-target comparison and is not assigned to a deletion analysis. No statistically established beyond-HSPA5 or candidate-specific advantage is claimed.

S5.6 Provenance and auxiliary robustness records

Cross-study gene-level tables and the primary 20-candidate DC score table permit direct inspection of their reported summaries. Deletion-specific candidate scores have separate provenance limits. The T-cell gene-level statistics have now been verified directly (S4.6). The two alternative-context digital exports are identical aliases of the same cross-study joined table; neither is a deletion-specific score table.

Table S5g | Evidence traceability

Result	Retained evidence	Unavailable detail
DC cross-study concordance	Summary and gene-level tables; 27 common scored targets per subtype	No missing detail for these descriptive summaries
Primary tumor-cDC1 ranking	All 20 candidate scores, target counts and primary permutation result	Deletion-specific results are separate
Shared-target deletion sensitivity	Archived narrative ranks and remaining XBP1 target counts	Eight-target identities; full candidate deletion scores and denominators; deletion-specific P
T-cell HSPA5 contrasts	Three exact author-artifact rows, condition and target metadata, SE, P, BH and derived Wald intervals	No local refit; study-level four-donor design retained

K562 auxiliary target-set robustness summaries are retained in the public release. They cover 121 archived subsets per score channel and eight leave-one-out configurations, with and without CSF3R. The subsets overlap and are not independent biological replicates. Their ranges provide a secondary description of sensitivity to sample composition.

Table S5h | Archived K562 robustness ranks

Target-set channel	Subsets	Rank range	Leave-one-out configurations	Leave-one-out rank range
Qualified edges; including CSF3R	121	1 to 12	8	3 to 9
Qualified edges; excluding CSF3R	121	1 to 12	8	3 to 9
Raw edges; including CSF3R	121	1 to 10	8	2 to 5
Raw edges; excluding CSF3R	121	1 to 10	8	2 to 5

The summaries use overlapping subsets; complete auxiliary score summaries remain in the deposit.

S6. Strict cross-seed direction agreement

A selected edge is concordant only when all three seed means are strictly positive or all are strictly negative. Edges containing zero values are nonconcordant. All selected values were finite. Counts refer to regulator-target edges, not distinct targets or biological replicates. Normalization changes the selected edge set and admits more unstable directions; it does not change the underlying seed-specific sign of an individual edge.

Under an illustrative null of three independent fair nonzero signs, the chance of all-positive or all-negative signs is 2 x (1/2)^3 = 0.25. More generally it is p^3 + (1 - p)^3 for independent signs with positive probability p. Neither assumption is established for correlated fits to the same data, and zeros also occur here. The observed rates are therefore descriptive technical-stability summaries, not significance tests against a universal 25% baseline.

Table S6a | Pooled strict sign agreement

Score	Selection	Concordant / total	Rate	Edges containing zero
Raw	Exact 50	967/1000	96.70%	2
Column L1	Exact 50	464/1000	46.40%	263
Column percentile	Exact 50	549/1000	54.90%	152
Raw	Including ties	967/1000	96.70%	2
Column L1	Including ties	464/1000	46.40%	263
Column percentile	Including ties	564/1026	54.97%	156

Percentage conversion is presentation of the archived fractions. Exact selections contain 50 edges per candidate; percentile boundary ties expand the denominator to 1,026.

Table S6b | Candidate-level agreement: exact / including ties

Candidate	Raw	Column L1	Column percentile
AHCTF1	47/50 ; 47/50	23/50 ; 23/50	30/50 ; 31/51
BCL11A	49/50 ; 49/50	21/50 ; 21/50	25/50 ; 26/52
CEBPA	48/50 ; 48/50	20/50 ; 20/50	28/50 ; 30/56
CEBPZ	49/50 ; 49/50	27/50 ; 27/50	29/50 ; 29/50
DRAP1	50/50 ; 50/50	21/50 ; 21/50	28/50 ; 30/52
ELF1	44/50 ; 44/50	22/50 ; 22/50	26/50 ; 26/50
ETV6	49/50 ; 49/50	21/50 ; 21/50	21/50 ; 22/52
GTF2I	49/50 ; 49/50	23/50 ; 23/50	25/50 ; 25/50
KDM2A	50/50 ; 50/50	26/50 ; 26/50	33/50 ; 33/50
MAZ	49/50 ; 49/50	28/50 ; 28/50	33/50 ; 35/53
NFAT5	49/50 ; 49/50	26/50 ; 26/50	30/50 ; 31/51
NME2	47/50 ; 47/50	23/50 ; 23/50	23/50 ; 23/50
PA2G4	49/50 ; 49/50	22/50 ; 22/50	30/50 ; 31/53
PBX3	48/50 ; 48/50	24/50 ; 24/50	26/50 ; 26/50
RUNX2	49/50 ; 49/50	27/50 ; 27/50	32/50 ; 32/50
SMYD3	50/50 ; 50/50	17/50 ; 17/50	12/50 ; 12/50
XBP1	49/50 ; 49/50	20/50 ; 20/50	28/50 ; 28/50
YBX1	49/50 ; 49/50	19/50 ; 19/50	23/50 ; 23/50
ZBED5	48/50 ; 48/50	31/50 ; 31/50	41/50 ; 44/55
ZBTB20	45/50 ; 45/50	23/50 ; 23/50	26/50 ; 27/51

Each cell lists exact-selection count/denominator followed by inclusive-selection count/denominator. The complete digital table preserves every selected edge and its three signed seed values.

S7. Complete existing-matrix scoring sensitivity

S7.1 Definitions and common comparison background

The prespecified grid comprised 17 scoring methods, K = 20/50/100, the same 20 candidates and both existing GRNBoost2 score definitions. All methods use the same three signed 1,191-by-3,006 matrices. Original raw, L1 and percentile scores were first reproduced against every row of the 60,100-row archived candidate-target export. Alternatives are descriptive sensitivity analyses; none was chosen by optimizing the three external case responses.

Self-edge order. In the original methods, self-edges enter column totals/rank populations before removal from output lists. Pre-exclusion removes each matching input-target entry before computing that column's denominator or percentile population. The percentile denominator becomes the number of remaining competitors (1,190 for columns with a self input; otherwise 1,191).

Column floors. Columns with original total below a quantile threshold are excluded from selection and assigned score zero. The 1st, 5th and 10th percentiles among strictly positive totals are 0.001348404589, 0.043558454880 and 0.072434176157. These are distributional filters, not estimated numerical-error thresholds. All three thresholds are reported.

Entry filters. The nonzero rule removes an input-target entry if any of the three signed seed means is zero. The same-sign rule requires all three means strictly positive or strictly negative. Each filter is applied before column normalization; percentile ranks and denominators use only retained competitors. Empty columns and excluded entries score zero. A 100% agreement rate after a same-sign filter is imposed by that rule and is not independent validation.

Z scores and ties. $Z(i,j) = [D(i,j) - \text{mean}_k D(k,j)] / \text{SD}_k D(k,j)$, using population SD over all 1,191 inputs. Constant columns receive zero. Selection retains only positive scores. The alternative percentile tie rule uses descending L1 score and then original feature index; percentile values and inclusive boundary ties are unchanged.

Common reference. Spearman correlations include all 3,005 non-self targets for every method, including original or filter-imposed zero scores, using average-tie ranks. Reference scores are the median official raw importance or median importance divided by each fit's n_estimators. Up to K positive proposed targets are compared with the reference top-K set including boundary ties. If fewer positive targets remain, all are retained without padding. This occurs only for SMYD3 under Z scoring (nine positive targets at each K). Counts sum candidate-target edges across 20 candidates; repeated targets are not independent observations. No inferential P values are assigned to this technical grid.

Table S7a | Complete method key

ID	Method
M01	Raw
M02	L1
M03	Percentile
M04	L1, self before
M05	Percentile, self before
M06	L1, floor 1%
M07	Percentile, floor 1%
M08	L1, floor 5%
M09	Percentile, floor 5%
M10	L1, floor 10%
M11	Percentile, floor 10%
M12	L1, nonzero before
M13	Percentile, nonzero before
M14	L1, same sign before
M15	Percentile, same sign before
M16	Column Z
M17	Percentile, L1 ties

S7.2 All methods at K = 50

Table S7b | Coverage and agreement at K = 50

ID	Edges	Distinct targets	Same sign (%)	Zero edges	Overlap raw / norm	rho raw	rho norm
M01	1000	87	96.7	2	63 / 25	0.4833	0.4166
M02	1000	532	46.4	263	21 / 24	0.3933	0.3546
M03	1000	620	54.9	152	22 / 21	0.3857	0.3557
M04	1000	532	46.4	263	21 / 24	0.3932	0.3545
M05	1000	623	54.8	156	22 / 21	0.3857	0.3558
M06	1000	528	47.3	254	21 / 24	0.3855	0.3514
M07	1000	620	55.4	149	22 / 21	0.3824	0.3526
M08	1000	518	51.2	214	21 / 23	0.3724	0.3451
M09	1000	610	59.1	116	22 / 21	0.3725	0.3425
M10	1000	505	53.4	187	21 / 24	0.3711	0.3441
M11	1000	597	61.0	98	22 / 21	0.3727	0.3467
M12	1000	515	65.9	0	24 / 25	0.4057	0.3615
M13	1000	603	66.4	0	24 / 24	0.4062	0.3622
M14	1000	446	100.0	0	16 / 18	0.2953	0.2640
M15	1000	617	100.0	0	13 / 15	0.2990	0.2708
M16	959	585	55.4	149	20 / 18	0.0676	0.0599
M17	1000	621	54.5	156	22 / 21	0.3857	0.3557

The sign fraction uses the actual number of selected edges: 959 for Z scoring and 1,000 otherwise. rho is the median of 20 full-target Spearman correlations. Raw/norm refer to the two GRNBoost2 references. IDs are defined in S7a.

S7.3 Alternative K values with all thresholds retained

Table S7c(20) | Coverage and overlap at K = 20

ID	Edges	Distinct	Universal	Same sign (%)	Zero	Overlap raw / norm
M01	400	40	6	97.3	0	21 / 2
M02	400	265	0	35.8	135	4 / 6
M03	400	305	0	47.3	75	1 / 3
M04	400	264	0	35.5	136	4 / 5
M05	400	303	0	46.3	75	1 / 3
M06	400	264	0	36.8	132	4 / 6
M07	400	303	0	47.3	75	1 / 3
M08	400	260	0	42.0	115	4 / 6
M09	400	304	0	51.7	56	1 / 3
M10	400	255	0	46.3	96	4 / 6
M11	400	303	0	54.5	45	4 / 4
M12	400	254	0	58.3	0	3 / 5
M13	400	297	0	59.3	0	3 / 3
M14	400	231	0	100.0	0	1 / 4
M15	400	299	0	100.0	0	3 / 3
M16	389	286	0	45.8	75	3 / 4
M17	400	296	0	46.0	76	1 / 3

Universal means present in every candidate list. Full-target rho does not depend on K and is reported in S7b. Z scoring yields only nine positive targets for SMYD3; all other method/candidate lists reach K.

Table S7c(100) | Coverage and overlap at K = 100

ID	Edges	Distinct	Universal	Same sign (%)	Zero	Overlap raw / norm
M01	2000	201	29	94.0	17	203 / 116
M02	2000	885	0	54.4	411	69 / 65
M03	2000	1034	0	60.0	273	75 / 67
M04	2000	885	0	54.3	410	68 / 64
M05	2000	1029	0	59.9	275	74 / 66
M06	2000	878	0	55.2	396	69 / 65
M07	2000	1025	0	60.5	261	77 / 68
M08	2000	869	0	58.1	335	71 / 69
M09	2000	1005	0	62.7	218	78 / 69
M10	2000	867	0	59.7	291	74 / 72
M11	2000	981	0	64.0	188	84 / 75
M12	2000	893	0	69.1	0	79 / 70
M13	2000	995	0	69.8	0	84 / 71
M14	2000	729	0	100.0	0	78 / 73
M15	2000	983	0	100.0	0	95 / 80
M16	1909	957	0	59.4	263	69 / 61
M17	2000	1016	0	59.7	281	74 / 67

Universal means present in every candidate list. Full-target rho does not depend on K and is reported in S7b. Z scoring yields only nine positive targets for SMYD3; all other method/candidate lists reach K.

S7.4 Case positions across every scoring alternative

Positions below are descending within-candidate order among positive-scoring non-self targets after the stated tie rule. A star marks exact top-50 membership. A position inside a tie is a deterministic selection position, not a unique score rank. CEBPA-MPO remains in the percentile top 50 with L1 ties (position 38), nonzero prefiltering (48), self pre-exclusion (35), and all tested column floors. Same-sign filtering excludes it because its three model signs are mixed. Z scoring also selects it (32) but has low all-candidate reference correlation (S7b).

Table S7d | Pair-specific target-list positions

ID	CEBPA-CSF3R	CEBPA-MPO	XBP1-HSPA5	MPO score-rank tier
M01	14 *	176	8 *	176
M02	419	190	733	190
M03	310	49 *	706	1-56
M04	419	190	734	190
M05	310	35 *	706	1-39
M06	418	189	733	189
M07	309	49 *	706	1-56
M08	403	180	718	180
M09	294	45 *	694	1-51
M10	383	176	699	176
M11	278	42 *	676	1-48
M12	373	160	673	160
M13	287	48 *	655	1-53
M14	479	Excluded	838	Filtered
M15	193	Excluded	515	Filtered
M16	421	32 *	712	32
M17	373	38 *	708	1-56

S8. Actual supervision and lineage-order audit

S8.1 Stored RNA-velocity supervision

The actual adata_rna.h5ad velocity layer was aligned to the matrix's 3,006 gene identifiers and the 4,828 training cells. There were 598 fully finite training columns, 2,408 entirely unavailable columns and no partially finite columns. The same columns were finite across all 6,035 cells. Ten of 13 universal raw targets were directly supervised. The direct RNA-velocity loss masks unavailable entries, while RNA-expression reconstruction and the cross-modal velocity-consistency term still constrain the outputs. Absence of direct RNA-velocity supervision is therefore a precise annotation, not evidence that the whole gene output was untrained.

Table S8a | Case target supervision and column scale

Target	Dynamics fit attempted	Finite training cells	Direct supervision	Column total
CSF3R	Yes	4828	Yes	2.2029406
MPO	Yes	0	No	0.5506026
HSPA5	Yes	4828	Yes	3.2987776

All three targets were included in the 1,000-gene dynamics-fit attempt; MPO has no finite output velocity and is masked in the direct RNA-velocity loss.

Table S8b | Target composition of original top-50 lists

Score	Supervised distinct / all distinct	Supervised edges / 1,000	Masked edges / 1,000
Raw	58 / 87	730	270
L1	98 / 532	183	817
Percentile	128 / 620	202	798

The universe is 598 supervised (19.9%) and 2,408 masked (80.1%) targets. Edge fractions measure repeated selections; distinct-target fractions measure unique genes.

Table S8c | Sign and zero rates stratified by supervision

Score	Direct velocity	Edges	Same sign (%)	Any zero (%)
Raw	Yes	730	97.7	0.3
Raw	No	270	94.1	0.0
L1	No	817	43.8	25.1
L1	Yes	183	57.9	31.7
Percentile	No	798	52.0	13.8
Percentile	Yes	202	66.3	20.8

The higher zero fraction is not confined to masked targets. Both supervision groups lose strict sign agreement under relative scoring. These strata are descriptive and share model seeds.

Table S8d | Full-universe column totals by supervision

Direct velocity	Targets	10th percentile	Median	90th percentile
No	2408	0E-7	0.1451302	0.4933575
Yes	598	0.0005313	0.1795071	0.9618285

S8.2 Existing lineage-order summaries

The archived model-time comparison supplies group medians from each selected seed. All three seeds order HSC before progenitor DC before DC. ProgDC-minus-HSC below is obtained by subtracting the two existing DC-referenced median differences. The GMP-to-granulocyte median difference is negative in every seed. Group-level ordering describes the sampled branch and does not identify a gene-specific dynamical trajectory or prove a derivative's accuracy.

Table S8e | Existing median model-time differences

Seed	ProgDC - HSC	DC - ProgDC	DC - HSC	Granulocyte - GMP
0	0.023142	0.205880	0.229022	-0.055062
1	0.038265	0.181551	0.219815	-0.078846
2	0.038446	0.269078	0.307525	-0.061257

Model-time units; the source is all_eight_priors_each_seed.tsv in the archived selected-checkpoint comparison. No trajectory or model was recomputed.

Supplementary data access and reading guide

Model weights, full matrices and source results are available at https://doi.org/10.5281/zenodo.24437365. The analysis code is available at https://github.com/dyc-max/dynavelo-target-scale-validation/releases/tag/v1.1.0. The Zenodo file 07_source_tables_and_scripts.zip contains the figure-source tables and analysis scripts. The Section S11 reproducibility archive supplies replay scripts and exact tables. Tables S7a–d report the scoring grid and case ranks, and Tables S8a–e report supervision and lineage-order results. Table S4f records the dendritic- and T-cell contrasts and identifies the published T-cell data source. Sections S9–S11 report the candidate-ranking intersection, independent evidence for ZBTB20–CENPF, and the perturbation benchmarks, overlap tests and cell-cycle sensitivity analysis.

Figure S2. Additional scoring and supervision comparisons

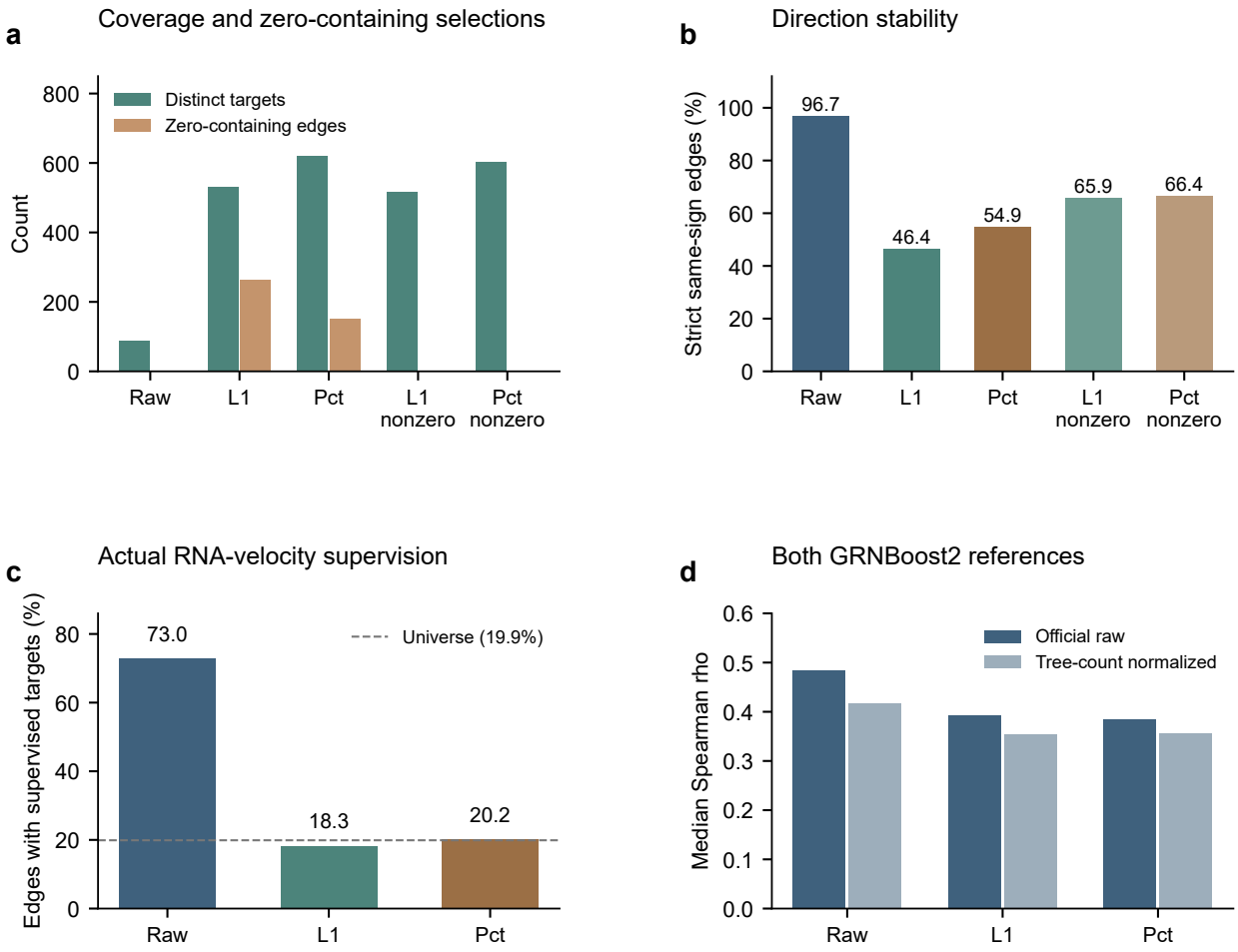


Figure S2. Panels a and b include nonzero-prefiltered scores alongside the three primary scores and report target coverage, zero-containing edges and sign agreement. Panel c compares the proportion of selected edges with direct velocity supervision to the supervised fraction of the target universe. Panel d compares rank concordance with two GRNBoost2 importance definitions on all 3,005 non-self targets. Sections S7 and S8 give all thresholds, strata and test settings.

S9. Candidate prioritization and curated reference lookup

The same 192-cell matrix and three weighted signed seed means used in Sections S1–S3 supplied the rankings and signs in Table 1. We ranked all 3,005 non-self targets within a regulator by descending score, with smaller feature index first for equal scores. For GRNBoost2 we used the median official raw importance across seeds. Cross-seed signs are the signs of the population-weighted DynaVelo seed means in the order 0, 1 and 2. Normalization changed the order in which edges were selected. The signed derivatives used here were computed in the frozen workflow.

The intersection required the first 50 positive-score positions by column L1 and GRNBoost2. Using the L1 order for this screen reduces the impact of large percentile ties at the cutoff. The percentile positions are reported alongside it. The shorter list required identical nonzero signs and absence of an exact directed human symbol pair in the archived CollecTRI v2.0 regulon file [18]. We matched all 43,175 rows of the release on 18 November 2026. The lookup records coverage of that file. This search did not systematically adjudicate the primary literature for all candidates.

Table S9a | All within-regulator column L1 and GRNBoost2 top-50 intersections

Regulator → target	Raw rank	L1 rank	Pct. rank	GRN rank	Seed signs	Shortlisted	CollecTRI v2.0
MAZ → HMGN5	324	11	25	11	− / − / −	Yes	No
GTF2I → S100A10	318	12	23	15	+ / − / +	No	No
MAZ → AC124248.1	948	17	105	12	+ / + / 0	No	No
ELF1 → WDR49	1102	6	1	28	− / − / −	Yes	No
DRAP1 → EVI2A	596	20	13	17	+ / + / +	Yes	No
SMYD3 → TPSB2	2079	17	25	26	− / − / +	No	No
XBP1 → HMGN5	363	14	34	33	− / − / −	Yes	No
YBX1 → UBE2S	318	29	20	21	− / − / −	Yes	No
CEBPA → KHDRBS3	663	20	134	34	− / − / −	Yes	No
PBX3 → E2F3	428	19	12	38	− / − / −	Yes	No
ETV6 → MYB	351	39	32	21	− / + / +	No	No
BCL11A → TANC2	600	27	37	34	+ / + / −	No	No
PA2G4 → ITPR2	67	50	21	13	+ / + / +	Yes	No
YBX1 → PRKN	235	16	14	49	+ / + / −	No	No
NFAT5 → MALAT1	41	43	23	28	− / − / −	Yes	No
ELF1 → FOXO3	535	48	31	26	+ / + / +	Yes	No
XBP1 → STAB1	400	32	59	50	− / − / −	Yes	No
BCL11A → SSBP2	296	46	28	38	+ / + / +	Yes	No
CEBPA → KHDRBS2	422	36	22	49	+ / + / −	No	No
CEBPA → ABHD2	928	43	561	45	+ / + / +	Yes	No
XBP1 → MEF2A	665	46	10	44	− / − / −	Yes	No

Rows are ordered by the sum of L1 and GRNBoost2 ranks, then regulator symbol and target feature index. Shortlisted rows have three identical nonzero signs and no exact directed symbol match in the archived CollecTRI file. The reference lookup was made on 18 November 2026. Database coverage is incomplete [18].

For the raw comparison, we ranked the 60,100 non-self edges across all 20 regulators separately by raw DynaVelo sensitivity magnitude and official raw GRNBoost2 importance, assigning tied scores their lowest shared rank number. ZBTB20–CENPF minimized the sum of these ranks (57 + 13 = 70). The respective within-regulator positions were 2 and 8. Its L1 and percentile positions were 388 and 368. The three signed means were +0.007773534, +0.009422577 and +0.007785908. The lack of a CollecTRI entry for this pair coexists with the external evidence in Section S10. ETV6–ZNF506 and PBX3–HOXA7 had L1 positions 14 and 17 but GRNBoost2 positions 55 and 146, respectively. Their seed signs were + / + / − and − / 0 / −.

S10. Independent evidence for ZBTB20–CENPF

S10.1 Thymic-organoid design and analysis

The mouse input and target orthologues are Zbtb20 (ENSMUSG00000022708) and Cenpf (ENSMUSG00000026605). We reanalyzed author-provided sparse count matrices from GSE267923 [19]. The study used CRISPR perturbations in murine hematopoietic stem cells and measured differentiated thymocytes. We compared sgZbtb20 with contemporaneous sgGFP controls. Filtered barcode suffixes were mapped to GEO libraries by matching 16-base cell barcodes to the prefix of public R1 reads, with 864–2,932 distinct matches in the assigned group and at most 11 matches in the next group. Batch 1 used the first guide and a CD3+/TCRβ+ sorting gate. Batches 2 and 3 used the second guide, with a CD4−/CD8− gate at week 2 and a CD3+/TCRβ+ CD4+/CD8+ selection at weeks 3–4. The source study reported altered coverage and splicing at the targeted Zbtb20 exon for the second guide, consistent with partial loss of function [19].

Table S10a | Public thymic-organoid libraries used for pseudobulk analysis

Batch	Week	Guide	GEO sample	Matrix group	Cells
1	2	sgGFP	GSM8281668	A:3	2,958
1	2	sgZbtb20-1	GSM8281670	A:5	965
1	3–4	sgGFP	GSM8281669	A:4	3,045
1	3–4	sgZbtb20-1	GSM8281671	A:6	6,849
2	2	sgGFP	GSM8765381	B:1	3,815
2	2	sgZbtb20-2	GSM8765383	B:3	3,019
2	3–4	sgGFP	GSM8765382	B:5	5,742
2	3–4	sgZbtb20-2	GSM8765384	B:7	4,854
3	2	sgGFP	GSM8765387	B:2	5,372
3	2	sgZbtb20-2	GSM8765389	B:4	3,797
3	3–4	sgGFP	GSM8765388	B:6	2,645
3	3–4	sgZbtb20-2	GSM8765390	B:8	5,046

A refers to the GSE267923 Aggregates matrix and B to the sgZbtb20_2 matrix. The number after the colon is the filtered barcode suffix. Each batch contributes a perturbation and a guide control at each time. The matched-protocol sensitivity analysis uses batches 2 and 3.

Analysis. We summed Gene Expression UMIs per library and filtered genes with edgeR filterByExpr using the relevant design. Models used TMM normalization, voom precision weights and robust empirical-Bayes moderation in limma [15, 16]. We fitted a batch term within each time point and used the perturbation coefficient for the contrast. Late Cenpf log2 CPM changes within the three batches were −0.413, −1.439 and −1.106 after TMM normalization. Whole-library contrasts include changes in the mix of sampled cells. Mki67 and Top2a also decreased in the late samples (all-batch log2 changes −0.865 and −0.795).

Table S10b | Cenpf response to Zbtb20 perturbation in mouse thymic organoids

Contrast	Batch pairs	log2 change	95% interval	Raw P	BH q	Genes tested
Week 2, all batches	3	−0.042	−0.185 to +0.102	0.516	0.806	15,303
Weeks 3–4, all batches	3	−0.952	−1.531 to −0.374	0.00914	0.338	15,509
Week 2, matched protocol	2	−0.001	−0.091 to +0.088	0.979	0.991	15,107
Weeks 3–4, matched protocol	2	−1.287	−1.510 to −1.063	5.32×10^{-6}	0.0037	14,845

Effects are sgZbtb20 minus sgGFP. The BH q values are adjusted across expressed genes in each contrast. Holm correction across the two time points for the nominated Cenpf target in all three batches gives P = 0.0183 at weeks 3–4 and P = 0.516 at week 2. The matched-protocol groups form a subset of the full analysis. Intervals are pointwise.

S10.2 Promoter occupancy and additional perturbation contexts

We intersected the human RefSeq Select/MANE CENPF transcription start site (NM_016343.4, GRCh38 chr1:214,603,195, 1-based) with the ENCODE conservative peak file [20]. The peak interval is chr1:214,602,767–214,603,266 (1-based inclusive), with its summit 178 bp upstream. The conservative set is derived from both replicates. The individual thresholded peak lists contained a TSS-overlapping peak for replicate 1 (ENCFF060VCF), with no overlap in replicate 2 (ENCFF693DXD). The binding assay used HepG2 cells. Table S10c gives the direction of responses in other public expression data. For SW620 we modeled the human RNA-seq counts with TMM–voom–limma, with two samples per hairpin and two shared scramble controls [22]. The combined contrast weights the two hairpins equally. In mouse CD8 T cells, each genotype and time point has one public aggregate library (GSM7471123–GSM7471126) [21]. The log2 CPM ratios from these libraries are descriptive. The plasma-cell comparison used robust, trend-adjusted limma on the deposited RMA-normalized array expression (Cenpf probe set 10360985, GPL6246) [23].

Table S10c | Additional ZBTB20–CENPF perturbation expression contrasts

Data and contrast	Replication	log2 change	95% interval	Raw P	BH q
SW620 shRNA 1	2 per group	−0.253	−0.452 to −0.054	0.0172	0.0623
SW620 shRNA 2	2 per group	+0.070	−0.127 to +0.267	0.449	0.604
SW620 mean of both hairpins	3 groups × 2	−0.091	−0.262 to +0.080	0.263	0.408
Mouse CD8, day 10	1 per genotype	−1.128	—	—	—
Mouse CD8, day 30	1 per genotype	−0.395	—	—	—
Mouse bone marrow plasma cells	4 per genotype	+0.518	+0.038 to +0.998	0.0371	0.755

Log2 expression differences are perturbation minus matched control, or log2 CPM ratios for the two aggregate CD8 comparisons. The combined SW620 estimate uses all six libraries. Intervals and P values are unavailable for descriptive library ratios.

S11. Benchmark diagnostics for target scale, context and cell-cycle sensitivity

S11.1 Design and scope

For each of the 20 regulators, four top-50 lists were fixed on the 3,005 non-self targets before the new perturbation outcomes were extracted. We used raw sensitivity, column L1, column percentile and the sum of within-regulator raw and official GRNBoost2 positional ranks. Equal scores were resolved by frozen target index. We retained the original lists when intersecting experimental coverage. The earlier XBP1–HSPA5 readouts and simple overlap baseline were already known when this retrospective benchmark was specified.

Primary screens comprised Replogle genome-wide K562 CRISPRi [17, 24] and Zhu primary CD4+ T cells at rest and 8 or 48 h after stimulation [12]. We used published gene-level effects and adjusted P values for Zhu. For Replogle we applied BH to all finite P values for each perturbation in the deposited differential-expression tables [24]. The main comparison required a decrease in the perturbed gene itself at gene $q < 0.05$ (8, 9, 11 and 11 regulators, respectively). Each target needed a measured effect and gene q. Missing readouts were recorded as unavailable. The directional endpoint also required identical, nonzero signs across the three seeds and a significant response opposite to the derivative sign after knockdown.

We permuted target identities jointly across regulators within discovery-expression and detection quintiles and identical evaluability patterns. This preserved shared-target structure. One-sided P values used (1 + exceedances)/(1 + 20,000). Benchmark BH families contained 32 primary or 16 secondary tests (four lists and two endpoints per environment). An exploratory fixed-target test reassigned pairs among targets with the same selection frequency and evaluability, preserving each regulator's list size.

S11.2 Target composition and response enrichment

Table S11a | Response enrichment in the Replogle CRISPRi screens

Screen	List	Hits / tested	Expected hits	Matched q	Assay-matched q	Fixed-target P
K562 genome-wide	Raw	17/188	5.60	0.0024	0.0156	0.801
K562 genome-wide	L1	7/187	4.57	0.621	0.465	0.24
K562 genome-wide	Percentile	10/197	4.84	0.202	0.0969	0.103
K562 genome-wide	Joint	21/268	8.51	0.0024	0.0246	0.839
K562 essential	Raw	8/78	2.21	0.008	0.037	0.252
K562 essential	L1	2/93	1.74	0.803	0.834	0.774
K562 essential	Percentile	4/102	2.00	0.265	0.198	0.235
K562 essential	Joint	16/141	4.45	0.0004	0.0024	0.604
RPE1 essential	Raw	20/122	9.56	0.0309	0.0969	1.000
RPE1 essential	L1	5/113	6.63	1.000	0.834	0.916
RPE1 essential	Percentile	4/120	7.21	1.000	1.000	0.961
RPE1 essential	Joint	43/175	13.64	0.0004	0.0012	0.198

The denominator counts evaluable regulator–target pairs. A hit is a transcriptome-adjusted target response (gene $q < 0.05$). Expected counts use discovery-sample matching. Assay matching uses expression and detection in 512 non-targeting control cells per screen (24 tests). Fixed-target P values are one-sided assignment tests. The K562 screens share regulators.

In genome-wide K562, raw and joint lists remained enriched under assay matching (7.51 and 11.68 hits expected). CEBPZ and DRAP1 accounted for 14/17 raw hits and 15/21 joint hits. The fixed-target control gave 17.52 and 22.85 expected, respectively. Its constraints allowed 84/188 raw and 256/268 joint assignments to vary. The data are consistent with selection of broadly responsive targets. The association between an individual target's column total and its experimental hit rate was not measured.

S11.3 Direction across cellular contexts

Table S11b | Significant responses concordant with the predicted direction

Screen	Regulators	Raw	L1	Percentile	Joint
K562 genome-wide	8	7/182 (0.202)	0/100 (1.000)	1/128 (1.000)	6/248 (0.621)
CD4+ T, resting	9	2/250 (1.000)	2/116 (0.734)	3/146 (0.7)	1/276 (1.000)
CD4+ T, 8 h	11	0/294 (1.000)	1/148 (0.736)	1/179 (0.753)	2/327 (0.736)
CD4+ T, 48 h	11	0/308 (1.000)	1/161 (0.736)	1/200 (0.736)	0/341 (1.000)
K562 essential	4	4/76 (0.0593)	0/45 (1.000)	1/62 (0.803)	7/129 (0.019)
RPE1 essential	5	9/120 (0.253)	0/62 (1.000)	0/80 (1.000)	7/161 (0.671)

Each entry gives concordant significant responses / evaluable predictions with stable signs (benchmark BH q). The K562 genome-wide screen and three T-cell conditions are primary. The two essential-gene screens form the secondary family. All four lists were assessed in RPE1.

In resting, 8-h and 48-h CD4+ T cells, the smallest q across either endpoint and all four lists was 0.383. In K562 essential, directional enrichment for the joint list (7/129, q = 0.0190) was sensitive to the matching scheme (assay-matched q = 0.0969). RPE1 joint-list estimates were 7/161, q = 0.671 under discovery matching and q = 0.834 under assay matching. Across primary screens, none of the methods showed directional enrichment at $q < 0.05$.

The discovery population was day-7 human bone-marrow CD34+ progenitors. K562 is a leukemia line, primary CD4+ T cells are mature lymphoid cells, and RPE1 is epithelial. Xbp1 deletion raised Hspa5 in mouse splenic and tumor dendritic cells (+0.378 and +0.413), whereas XBP1 CRISPRi reduced HSPA5 in stimulated human T cells (−0.818). Species, perturbation modality and cell state also differed (Section 2.3). These comparisons leave open how accurately signs transfer within the progenitor population. The number of confirmed perturbations and significant sign-concordant observations limits the precision of the comparison.

S11.4 Top-50 intersection and continuous concordance

Across 20 regulators, 21 edges entered both column L1 and official GRNBoost2 top-50 lists (Table S9a). Independent uniform lists yielded 16.639 expected (one-sided P = 0.167, convolution of 20 hypergeometric distributions). Reassigning the 20 intact GRNBoost2 lists to regulators while excluding self-edges gave 17.754 expected across 100,000 accepted permutations (95% null interval 9–28, P = 0.267). Among sign-stable edges, 14 were observed and 11.046 expected (P = 0.240). The raw lists had 63 overlaps with GRNBoost2, compared with 21 for L1 and 22 for percentile (Table S1a).

Over all 3,005 non-self targets per regulator, median Spearman correlations with official GRNBoost2 were 0.4833 for raw, 0.3933 for L1 and 0.3857 for percentile. Across the 20 regulators the minima were 0.285 for L1 and 0.287 for percentile (Tables S1a,b). Excluding zero-containing entries before normalization gave medians of 0.4057 and 0.4062 (Table S7b). Both algorithms used the discovery data. Their positive rank correlations coexist with a top-50 intersection compatible with the random baseline.

Of the 21 intersection edges, 13 had usable external readouts in at least one primary environment, nine also met the knockdown criterion, and seven had stable signs. Across those seven edges, 0/20 evaluable edge–condition observations were both significant and directionally concordant. The sole significant response was YBX1–UBE2S at 48 h in CD4+ T cells (log2 change −0.548, gene q = 0.0457). The three negative model derivatives predicted an increase after knockdown. Edges without adequate coverage remain unresolved.

S11.5 ZBTB20–CENPF and cell-cycle composition

The 12 GSE267923 libraries from three paired batches contained 48,107 cells, including 28,181 at weeks 3–4 [19]. We scored G1-like, S and G2/M phases with the Seurat 2019 marker sets and expression-matched controls. Cenpf was excluded from the signatures, control genes and scoring-library denominator. Phase counts summed exactly to the archived gene counts for each library. We fit the batch-paired TMM–voom–limma models described in Section S10 within each time and phase, requiring at least 20 cells per arm per participating batch.

Table S11c | Late thymocyte Cenpf responses within cell-cycle phases

Comparison	log2 change	95% interval	Raw P	Holm P	Gene BH q
All cells, three batches	−0.952	−1.531 to −0.374	0.00914	—	0.338
G1-like, three batches	−0.815	−1.430 to −0.200	0.0161	0.0805	0.389
S, three batches	−0.529	−0.960 to −0.097	0.0205	0.082	0.642
G2/M, three batches	−0.265	−0.464 to −0.065	0.0126	0.0758	0.422
G1-like, batches 2–3	−1.271	−1.779 to −0.762	0.000378	0.00227	0.0367

Within-phase estimates compare sgZbtb20 with sgGFP in batch-paired pseudobulks. Intervals are pointwise. Holm adjusts the six Cenpf tests across two times and three phases, separately for the full and matched-protocol analyses. BH adjusts across expressed genes within each contrast.

All three late phase-specific estimates were negative, and the three-batch Holm P values ranged from 0.0758 to 0.0820. The G1-like estimate was largest in magnitude. In batches 2 and 3, which shared a guide and a CD4+CD8+ sorting gate, that estimate was −1.271 (gene q = 0.0367). This two-batch subset was reported previously for its shared protocol (Section S10) and examined separately as a sensitivity analysis of the full three-batch comparison.

For direct standardization we calculated Cenpf per 10,000 RNA UMIs in each cell and gave the two arms the same phase proportions within each batch, using the mean of their observed proportions. The equally weighted mean of the three batch log2 ratios changed from −1.071 before standardization to −0.526 after standardization (paired 95% t interval −2.079 to +1.026, two-sided P = 0.282). The adjusted batch estimates were +0.112, −0.554 and −1.137. Further stratification by Rag1/Rag2 and Itm2a detection gave −0.469 (95% interval −1.808 to +0.870). At week 2, phase standardization gave +0.134 (−0.115 to +0.383). The all-cell model in Table S11c uses library-level counts, so attenuation is assessed using the two direct-standardization estimates on the same scale.

S11.6 Scope of inference

The screens support enrichment of responsive targets in defined settings. Regulator-specific assignment and sign accuracy in matched progenitors remain open. The 14 sign-stable intersection edges are computational candidates. ZBTB20–CENPF was selected by the global raw-plus-GRNBoost2 rank sum (within-regulator L1 position 388). Its reduction in the matched-protocol G1-like subset and the wider standardized interval across three batches motivate a cell-state-matched perturbation test. The extent of any direct regulatory contribution remains unresolved.