Cross-cell transfer of reusable chemical-response programs in Z-Screen

Abstract

Multi-context cellular screening is powerful but scales poorly when every compound must be profiled in every cell type. Z-Screen addresses this cost by linking one-bead-one-compound chemistry, building-block provenance, tuple identity, imaging-derived features, and low-pass RNA state in a shared discovery system. Here we test whether chemical responses measured deeply in one cell context can prioritize where to measure next. In the canonical Z-Screen public dataset, two paired library/cell-line systems support primary molecule-level transfer: ZEL024 H1650/HEK293 and ZEL031 A549/THP1. At the molecule level, ridge transfer made the improvement explicit: in ZEL024, centered-response cosine rose from 0.146 with direct source reuse to 0.371 for H1650->HEK293 and 0.279 for HEK293->H1650; in ZEL031, it rose from 0.094 to 0.242 for A549->THP1 and 0.212 for THP1->A549. The strongest result emerged when molecules were grouped into chemistry-resolved response programs: transfer increased support-weighted cosine from a 0.471 identity baseline to 0.808-0.863 for ZEL024 bb0+bb1 programs, and from 0.419 to 0.664-0.714 for ZEL031 bb0 programs. ZEL028-2 single-building-block aggregates provided useful guardrails, showing that high scores must preserve chemical specificity. Cross-library projection and gene-level decoding produced directional, hypothesis-generating signals and define the next frontier. These results establish cross-cell transfer as a practical Z-Screen scaling strategy: profile combinatorial chemistry deeply in one context, learn reusable chemical-response programs, and prioritize target-cell follow-up with explicit uncertainty.

Significance

Adding cell types is one of the fastest ways to increase both the value and the cost of a functional chemistry screen. Z-Screen creates a practical alternative to measuring every compound everywhere. The platform can learn how chemistry-linked RNA programs move between paired cell contexts, then use those maps to rank which target-cell measurements are most worth performing. The most transferable unit in the current public data is not the isolated sparse molecule; it is the reusable chemical-response program defined by shared combinatorial building blocks at a level that remains chemically specific. That is the level at which early discovery teams often make series-level decisions.

Introduction

Cell type is not a nuisance variable in chemical biology. It determines pathway wiring, target abundance, stress response, lineage state, transport, and the transcriptional consequences of perturbation. A compound series that looks promising in one cell type may become inactive, toxic, or mechanistically different in another. The direct answer is to run every compound across every relevant cell context. That strategy is scientifically clean, but it scales poorly as chemistry, dose, timing, and cellular diversity expand.

Z-Screen is designed for a complementary operating model. It measures large one-bead-one-compound libraries while preserving the chemical provenance of each profiled molecule. A building block is a synthetic component used to assemble combinatorial molecules; a tuple is the ordered combination of building blocks that defines a compound in a library. In the public bundle, raw Z-Screen SMILES are replaced with irreversible `smiles_hash` identifiers, and chemistry-aware analyses use building-block annotations plus precomputed chemistry embeddings. The platform advantage is that compound identity, tuple structure, RNA state, and imaging-derived measurements remain linked, allowing chemical-response programs to be studied as reusable objects rather than as isolated screening hits.

This manuscript asks whether that structure can reduce the cost of multi-context biology. We define cross-cell transfer as a measured source-cell response, a paired landmark set of compounds observed in both source and target cell lines, and a learned map that estimates or prioritizes the target-cell response for held-out chemistry. The goal is not to replace direct target-cell validation. The goal is more useful and more immediate: profile chemistry deeply in one context, learn how response programs translate across biology, and decide which target-cell experiments deserve the next measurements.

Public perturbational atlases provide the broader field context. LINCS L1000 established large-scale transcriptional perturbation mapping in a reduced 978-gene representation [2]. Tahoe-100M extends single-cell drug perturbation profiling across 50 cancer cell lines for roughly 1,100 to 1,200 compounds [3]. scGeneScope contributes treatment-matched imaging and transcriptomics as a multimodal benchmark [4]. Z-Screen occupies a different axis: one of the largest public combinatorial chemistry transcriptomic datasets, with supporting imaging-derived features and explicit building-block and tuple provenance. In the full public bundle, Z-Screen contains 615,793 repaired RNA profiles, 615,721 valid scVI latent profiles, 12 combinatorial libraries, 4 cell lines, and 142,187 unique hashed compounds with chemistry embeddings. The relevant question for this paper is whether those chemistry coordinates make cross-cell response transfer useful.

We separate three levels of evidence. First, held-out molecule-level transfer tests whether a source-cell signature carries target-cell information beyond direct reuse. Second, grouped-program transfer asks whether molecules sharing defined building-block programs form more stable transferable response units. This is the primary platform result. Third, cross-library projection and gene-level decoding ask whether transfer can be extended beyond the current dense paired library settings. These frontier analyses are reported because they define where the platform can scale next.

Results

Paired library coverage defines the transfer opportunities

Question. Which library/cell-line pairs provide enough shared chemistry to test cross-cell transfer?

Comparison. The canonical Z-Screen public dataset is broad but uneven across libraries and cell lines. ZEL024 contains H1650 and HEK293 runs with 10,695 and 13,923 unique compounds, respectively; HEK293 is the denser side, with a median of 11 wells per compound. ZEL031 contains A549 and THP1 runs with 8,321 and 9,041 unique compounds, plus a much smaller H1650 run. ZEL028-2 is broad and shallow, with 40,622 A549, 25,906 H1650, and 61,396 HEK293 unique compounds, each at a median of 1 well per compound.

Metric. Eligibility was based on shared compound count after a minimum per-cell support filter. At a two-cell filter, ZEL024 H1650/HEK293 retained 6,299 shared compounds, and ZEL031 A549/THP1 retained 3,561 shared compounds. ZEL028-2 overlaps fell to 9-19 shared compounds at the same filter, and ZEL031 comparisons involving H1650 were similarly too small. At a five-cell filter, the eligible molecule-level set narrowed to 353 shared compounds for ZEL024 H1650/HEK293 and 140 for ZEL031 A549/THP1.

Interpretation. These coverage facts set the evidence hierarchy. ZEL024 and ZEL031 support primary held-out molecule-level transfer. The broader ZEL028-2 data are most informative as grouped-transfer diagnostics, aggregation guardrails, and cross-library stress tests. This structure is exactly the scaling problem Z-Screen is built to address: deep paired maps in some contexts can guide which additional cell-type measurements should be prioritized.

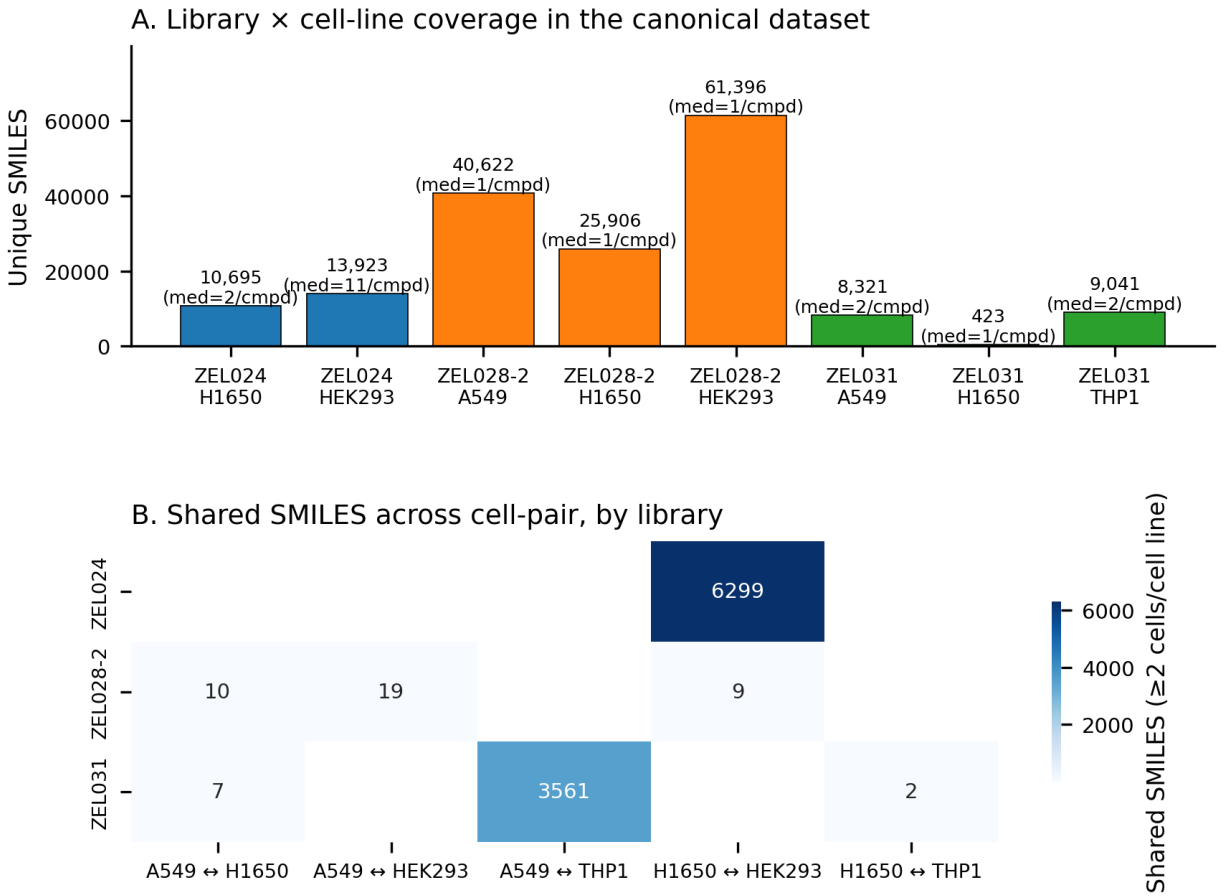


Figure 1. Library-by-cell-line coverage in the canonical Z-Screen dataset. Top: unique compounds and median wells per compound per system. Bottom: shared-compound counts across cell-pair directions, by library. The two dense paired systems used in the molecule-level transfer benchmark are ZEL024 (H1650 <-> HEK293, 6,299 shared) and ZEL031 (A549 <-> THP1, 3,561 shared).

Learned maps improve held-out molecule-level transfer

Question. For individual compounds measured in two cell lines, does a learned transfer map recover target-cell response better than simply reusing the source-cell signature?

Comparison. Each compound was represented by the mean of its 32-dimensional scVI latent coordinates within a library and cell line. The benchmark used held-out compounds from the two eligible paired systems: ZEL024 H1650 <-> HEK293 and ZEL031 A549 <-> THP1. Direct source-signature reuse, labeled identity, was compared with ridge transfer, partial least squares (PLS), and k-nearest-neighbor transfer.

Metric. The primary metric was held-out row-wise cosine similarity in centered-response space. Centering subtracts source and target training means before fitting and evaluation, focusing the benchmark on perturbation-linked variation rather than constant cell-line baseline shifts.

At the largest training sizes, learned maps consistently improved on identity:

Library	Direction	Train size	Identity	Ridge transfer	PLS transfer
ZEL024	H1650 -> HEK293	5,000	0.146	0.371	0.358
ZEL024	HEK293 -> H1650	5,000	0.146	0.279	0.275
ZEL031	A549 -> THP1	3,000	0.094	0.242	0.242
ZEL031	THP1 -> A549	3,000	0.094	0.212	0.214

Interpretation. Molecule-level transfer is detectable and useful as a prioritization signal. Direct reuse leaves substantial target-cell information uncaptured, while simple regularized maps recover additional response structure in every eligible direction. The absolute centered-response cosines also show why the molecule is not the strongest current unit of transfer: many compounds have limited well support, and low-pass transcriptomic measurements are noisy at single-molecule resolution. Z-Screen becomes more powerful when chemistry is organized at the program level.

Raw-profile benchmarks are retained as supplementary diagnostics because they capture baseline cell-state translation. In raw space, transfer scores are much higher; for example, ZEL024 HEK293 -> H1650 reaches 0.546 with ridge transfer at train size 5,000 versus 0.063 with direct reuse. The centered-response benchmark remains the main discovery readout because it better isolates perturbation-linked signal.

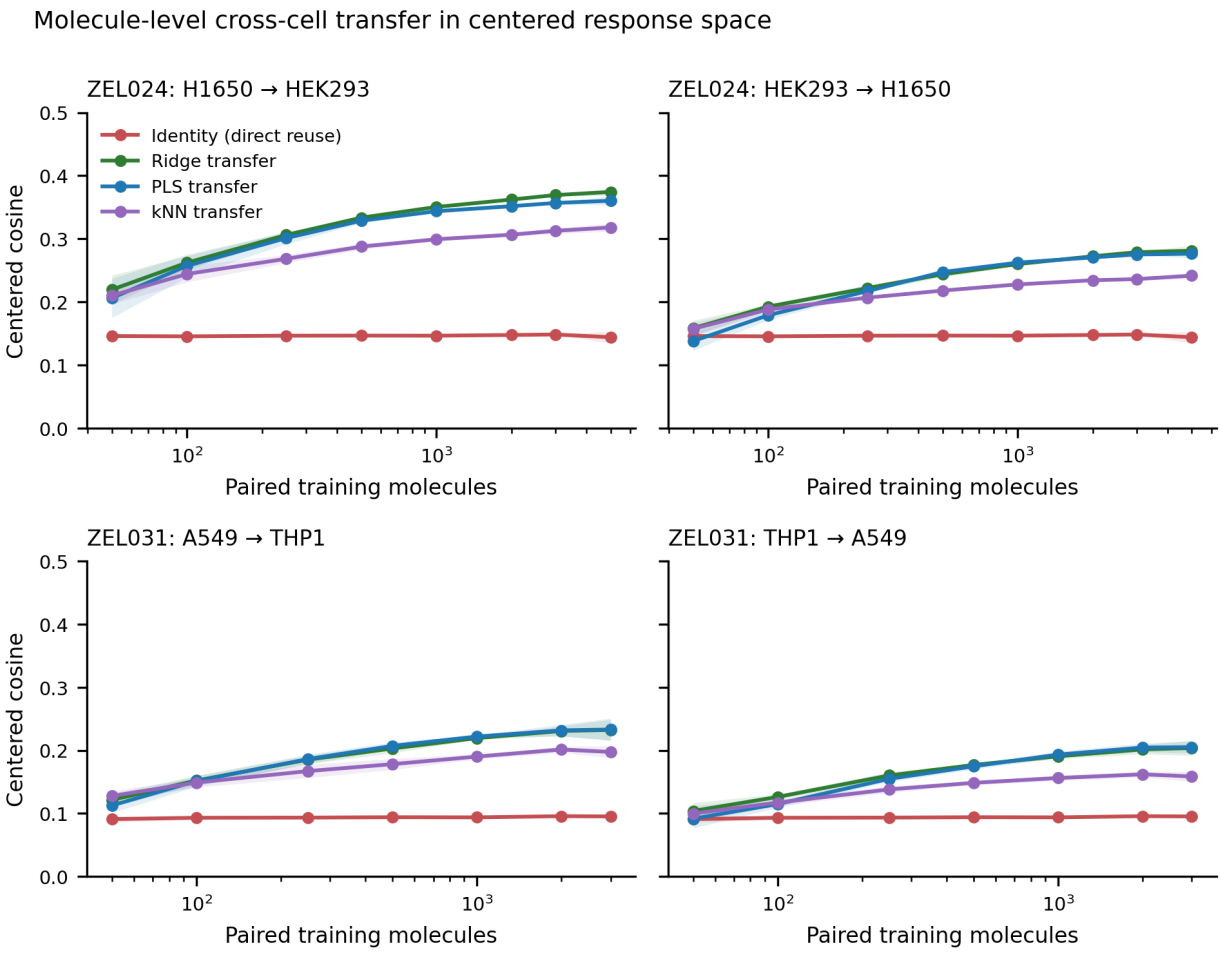


Figure 2. Molecule-level cross-cell transfer in centered-response space across four well-supported cell-pair directions in ZEL024 and ZEL031. Direct reuse of the source signature is consistently weaker than learned transfer; ridge and PLS transfer are nearly tied.

Chemistry-resolved response programs transfer strongly across cell types

Question. Are chemical-response programs defined by shared building-block provenance more reusable across cell types than individual molecules?

Comparison. We averaged molecules into grouped programs at aggregation levels that remain chemically interpretable. In ZEL024, which has four substantive building-block positions, the main program unit is a bb0+bb1 pair. In ZEL031, which has two substantive positions, the bb0 group is the closest comparable program because the other position remains the varying partner chemistry. These groupings preserve enough chemical specificity to support series-level decisions while increasing measurement support per response centroid.

Metric. The benchmark used centered ridge transfer and summarized held-out performance as support-weighted cosine similarity across shared programs. Support weighting gives greater influence to program centroids with more measured wells.

Library	Direc-tion	Group defini-tion	Shared groups	Median support	Identity	Ridge transfer	Gain
ZEL024	H1650 -> HEK293	bb0+bb1	84	240.5	0.471	0.863	0.392
ZEL024	HEK293 -> H1650	bb0+bb1	84	240.5	0.471	0.808	0.336
ZEL031	A549 -> THP1	bb0	129	157.0	0.419	0.714	0.296
ZEL031	THP1 -> A549	bb0	129	157.0	0.419	0.664	0.245

Interpretation. This is the central result. Chemistry-resolved grouped programs transferred with support-weighted cosine 0.664-0.863 across four cell-pair directions, improving over identity by 0.245-0.392 absolute cosine. The result turns cross-cell transfer from a molecule-by-molecule prediction exercise into a platform scaling strategy. A deep source-cell profile can reveal reusable response programs; paired landmark maps can translate those programs into target contexts; and follow-up screens can focus on the programs most likely to matter in the next cell type.

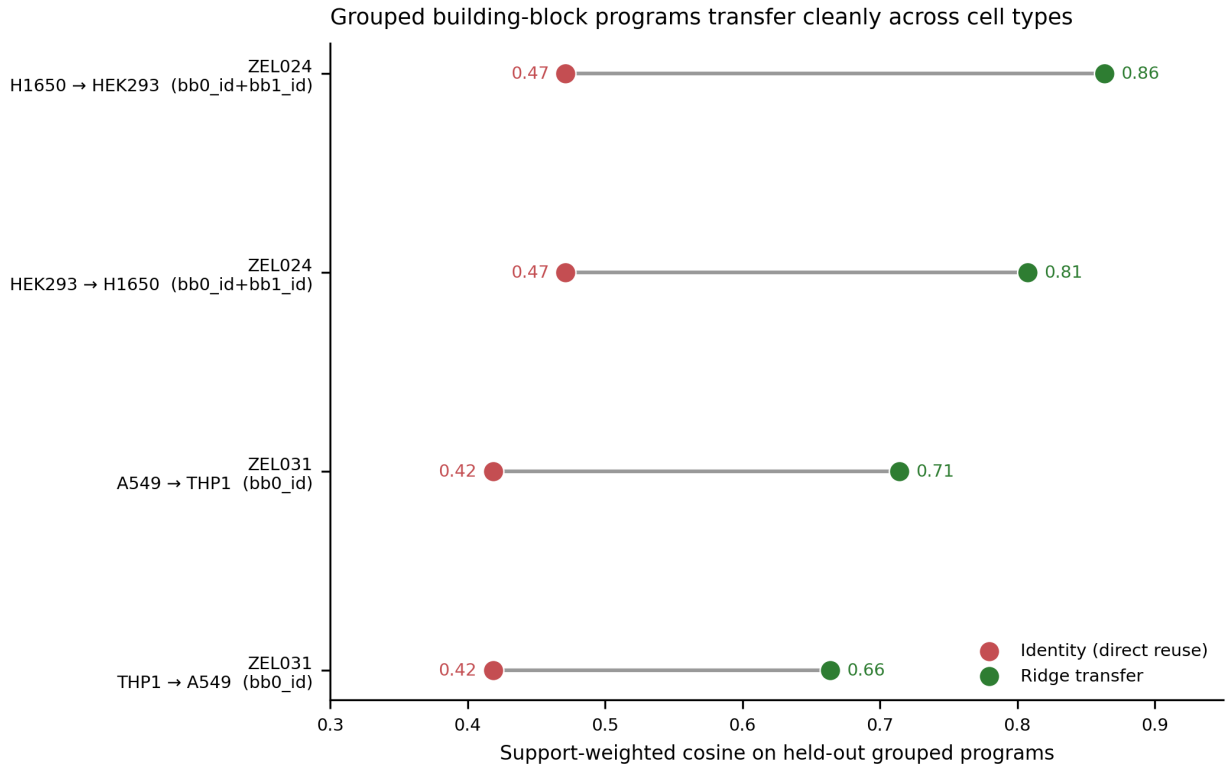


Figure 3. Grouped building-block program transfer at chemistry-resolved aggregation levels (bb0+bb1 for ZEL024, bb0 for the two-position ZEL031 library). Identity sits in the 0.42 to 0.47 range, while a regularized learned transfer map reaches support-weighted cosine 0.66 to 0.86 across four chemistry-resolved program-and-direction combinations.

The side-by-side summary highlights the scale advantage. At the molecule level, learned transfer roughly doubled centered-response cosine over identity in ZEL031 and improved ZEL024 by 0.13-0.23 absolute cosine. At the grouped-program level, learned transfer improved over identity by 0.25-0.39 absolute cosine, with all four reported directions ending in the 0.66-0.86 range.

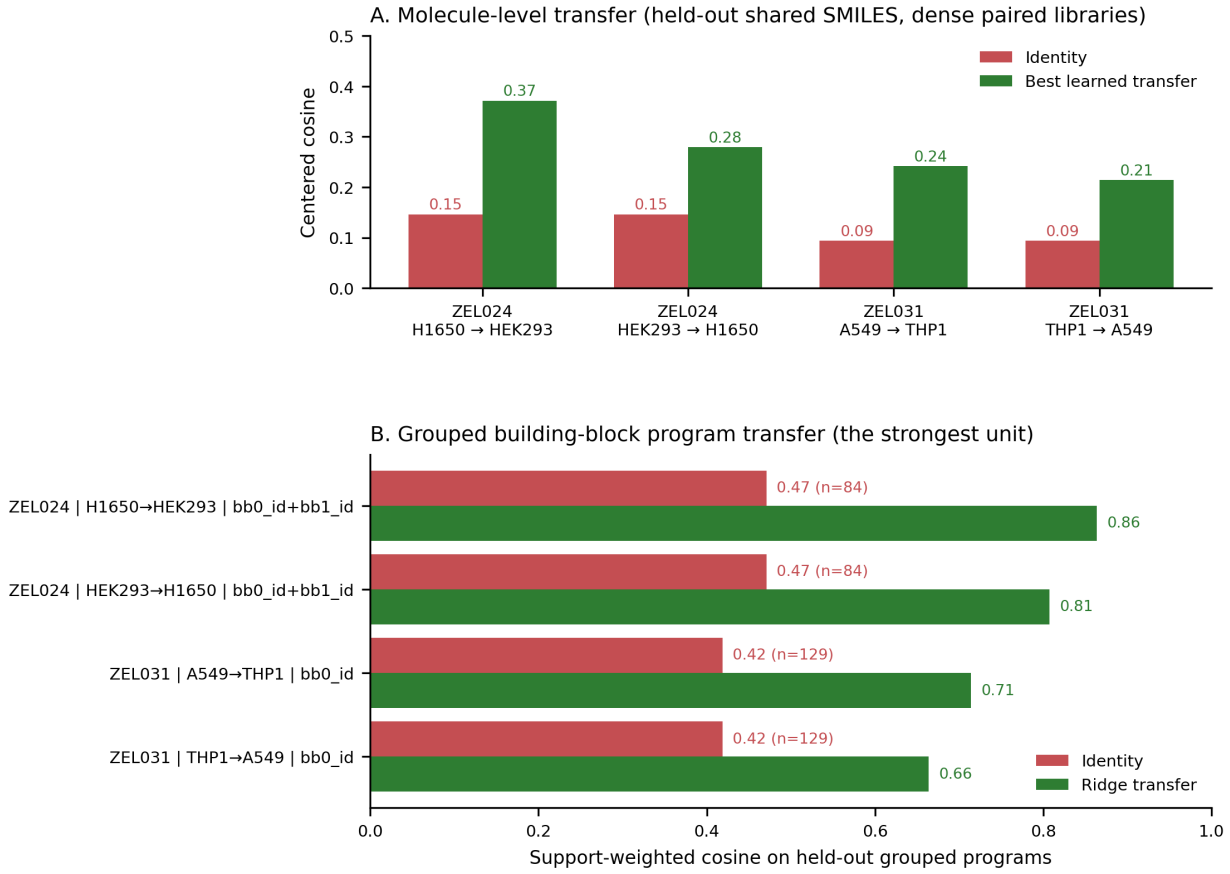


Figure 4. Best learned transfer versus identity at both molecule level (top) and grouped building-block program level (bottom). Learned transfer wins in every eligible cell-pair direction tested, with the largest gains at the grouped-program level.

Aggregation guardrails preserve chemical specificity

Question. When grouped-transfer scores are high, do they still represent chemically specific programs?

Comparison. ZEL028-2 provides an important guardrail. It has four substantive building-block positions, but some of its strongest grouped-transfer rows come from a single building-block position. A single-position centroid in this library averages over many distinct tuples and can report broad

library/cell-pair structure rather than a specific chemical-response program.

Metric. In the full grouped-transfer table, ZEL028-2 single-building-block rows reached high weighted cosines. For example, bb1 grouping reached 0.850 for A549 -> HEK293 and 0.837 for H1650 -> HEK293. More specific BB-pair rows in the same library were much shallower per group and scored lower.

Direction	Group	Class	Groups	Median support	Transfer	Gain	Interpretation
A549 -> HEK293	bb1_id	single-BB aggregate	86	526.5	0.850	0.332	high transfer but over-aggregated
H1650 -> HEK293	bb1_id	single-BB aggregate	86	333.5	0.837	0.366	high transfer but over-aggregated
HEK293 -> A549	bb1_id	single-BB aggregate	86	526.5	0.762	0.244	high transfer but over-aggregated
A549 -> HEK293	bb1_id+bb2_id	BB-pair group	6,690	5.0	0.227	0.100	more specific but shallow per group
A549 -> HEK293	bb3_id+bb4_id	BB-pair group	7,158	5.0	0.215	0.089	more specific but shallow per group
HEK293 -> A549	bb1_id+bb2_id	BB-pair group	6,690	5.0	0.213	0.086	more specific but shallow per group

Interpretation. The guardrail strengthens the grouped-program claim by defining what does and does not count as a reusable chemical-response program. High transfer is most meaningful when

the grouping remains chemically specific enough to guide synthesis or follow-up selection. ZEL024 bb0+bb1 and ZEL031 bb0 satisfy that standard in their library designs. ZEL028-2 single-building-block aggregates are useful diagnostics of transferable broad structure, but they are not the main platform evidence.

Cross-library projection marks the next scaling frontier

Question. Can a transfer map learned in one combinatorial library be applied to another library measured in the same cell-pair direction?

Comparison. A centered ridge transfer matrix was trained in one library and applied to compounds from another library. This setting is harder than within-library transfer because the held-out library can differ in building-block vocabulary, scaffold composition, sampling depth, and overlap structure.

Metric. Performance was summarized as support-weighted cosine in centered-response space, comparing identity with ridge transfer in the held-out library.

Projection	Identity	Ridge transfer	Gain
ZEL024 -> ZEL028-2, H1650 -> HEK293	0.033	0.055	0.022
ZEL024 -> ZEL028-2, HEK293 -> H1650	0.033	0.059	0.026
ZEL028-2 -> ZEL024, H1650 -> HEK293	0.119	0.121	0.002
ZEL028-2 -> ZEL024, HEK293 -> H1650	0.119	0.067	-0.052
ZEL028-2 -> ZEL031, A549 -> H1650	0.120	0.216	0.096
ZEL028-2 -> ZEL031, H1650 -> A549	0.120	0.139	0.019
ZEL031 -> ZEL028-2, A549 -> H1650	0.103	-0.009	-0.112
ZEL031 -> ZEL028-2, H1650 -> A549	0.103	0.024	-0.080

Interpretation. Cross-library projection is a frontier result. Some directions were positive, including both ZEL024 -> ZEL028-2 HEK293/H1650 projections and ZEL028-2 -> ZEL031 for A549 -> H1650. Other directions were weak or negative. The pattern shows that cross-cell maps can carry signal beyond a single library, but teacher-library depth, target-library support, and chemistry-domain match matter. The next prospective experiment should be designed explicitly for this question, with planned shared chemistry, balanced cell-line sampling, scaffold-family overlap, and predefined rejection criteria.

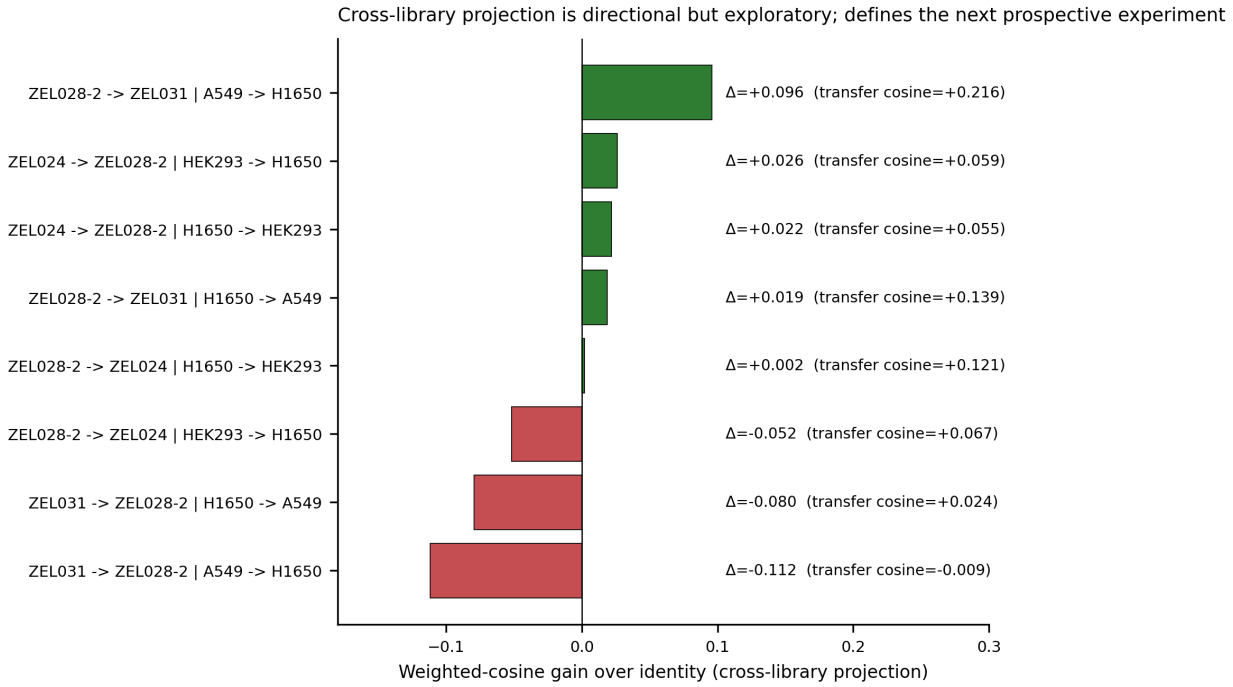


Figure 5. Cross-library projection results. Some cross-library projections are positive, but performance remains small, asymmetric, and sensitive to which library acts as teacher.

Gene-level decoding turns transferred latents into ranked hypotheses

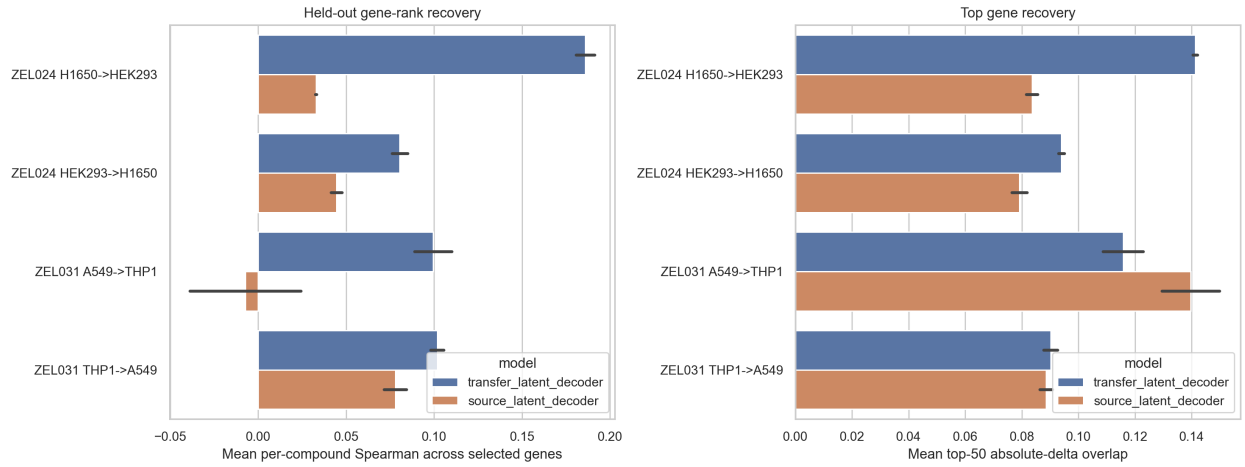
Question. Can transferred latent profiles be decoded into ranked gene-level hypotheses in the target cell type?

Comparison. This experiment composed two models: a source-to-target latent transfer map fit on shared compounds, and a target-cell ridge decoder fit from target-cell scVI centroids to log-CPM pseudobulk deltas over 800 target-training variable genes. A rank signature is the ordered list of genes most increased or decreased by a perturbation. The goal here was ranked gene-hypothesis generation, not full transcriptome reconstruction.

Metric. Held-out predictions were scored by per-compound Pearson and Spearman correlations over gene deltas and by top absolute-delta gene overlap. The clearest gain occurred in ZEL024 H1650 -> HEK293: the direct source-latent decoder reached mean gene-delta Spearman 0.033 and top-50 absolute-delta overlap 0.084, while the transferred-latent decoder reached Spearman 0.186 and top-50 overlap 0.141. An oracle decoder using the measured target latent reached Spearman 0.353 and top-50 overlap 0.187. The reverse ZEL024 direction and both ZEL031 directions produced smaller transferred-latent Spearman values of 0.081-0.102.

Interpretation. Gene-level decoding extends cross-cell transfer from latent prioritization toward interpretable biology. The current signal is sufficient to nominate ranked target-cell genes or pathways for inspection, especially in ZEL024 H1650 -> HEK293, while the oracle gap shows room for better target-cell decoders and richer training data. This is a natural next layer for Z-Screen:

transfer the program, decode the target-cell hypothesis, then measure the most decision-relevant cell type directly.



Supplementary Figure 2. Exploratory gene-level transfer. A target-cell decoder composed with the learned source-to-target latent map partially recovers held-out gene-rank structure, with the clearest gain in ZEL024 H1650 -> HEK293 and weaker gains in the remaining directions.

Discussion

Cross-cell transfer is a platform scaling advantage for Z-Screen. The strongest evidence is that chemistry-resolved response programs, not just individual compounds, can be translated across cell types with high support-weighted cosine. In ZEL024, bb0+bb1 programs transferred between H1650 and HEK293 at 0.808-0.863. In ZEL031, bb0 programs transferred between A549 and THP1 at 0.664-0.714. These are reusable chemical-response programs: they retain building-block specificity, average over enough measured chemistry to stabilize the RNA signal, and map onto the way discovery teams reason about series-level follow-up.

Molecule-level transfer provides a complementary result. Learned maps improved over direct source reuse in every eligible dense direction, with centered-response cosine reaching 0.279-0.371 in ZEL024 and 0.212-0.242 in ZEL031. These values are not the headline because sparse single-molecule profiles are noisy, but they matter operationally. They show that paired landmark data capture compound-specific target-cell information and can help rank molecules within or around a transferable program.

The ZEL028-2 guardrail clarifies the chemical standard for this claim. Single-building-block aggregates in a four-position library can score highly because they average broad library structure. Those rows are valuable diagnostics, but they are too coarse to act as the main chemical-program evidence. The stronger platform interpretation comes from groupings that preserve chemistry specificity while increasing support: bb0+bb1 in ZEL024 and bb0 in the two-position ZEL031 design.

Cross-library and gene-level analyses define the next build-out. Cross-library projection already shows directional gains, but performance depends on which library acts as teacher and how well the

chemistry domains overlap. Gene-level decoding shows that transferred latents can be converted into ranked target-cell hypotheses, with a clear ZEL024 H1650 -> HEK293 gain and measurable headroom relative to oracle target-latent decoding. Together, these analyses point to a prospective architecture in which Z-Screen builds a shared chemical-response latent space, learns cell-type-specific adapters, rejects out-of-domain projections, and allocates new target-cell measurements where uncertainty and expected information gain are highest.

The practical workflow is immediate. Profile a combinatorial library deeply in a source cell type. Use a paired landmark panel to learn how source responses translate into a target cell type. Identify chemical-response programs that are likely to retain, lose, invert, or shift activity across contexts. Then measure the most informative programs and representative molecules directly in the target cell line. This turns cross-cell biology from an exhaustive grid into an adaptive measurement problem.

Methods

Data inputs

All analyses used the canonical Z-Screen RNA aggregate dataset distributed with this repository:

- `data/ZScreen_Canonical_Dataset/RNASeqAggregate/canonical_obs.parquet`
- `data/ZScreen_Canonical_Dataset/RNASeqAggregate/canonical_scvi_latents.parquet`
- `data/ZScreen_Canonical_Dataset/RNASeqAggregate/canonical_chemistry.parquet`
- `data/ZScreen_Canonical_Dataset/RNASeqAggregate/chem_embed.parquet`

Rows were retained when they had a non-null `smiles_hash` and `chemistry_grain == "building_block_annotation"`. In the public bundle, `smiles_hash` is the chemistry-resolved compound identifier used for joins; raw Z-Screen SMILES are not included. Compound-control-only rows were excluded from the primary transfer benchmark because their annotations are mechanism-of-action labels rather than combinatorial tuple coordinates.

Latent representation

Transfer benchmarks used 32-dimensional scVI latent coordinates [1]. Each well in the RNA aggregate is a low-pass pseudobulk measurement from a small number of cells, so individual genes are noisy at the single-well level. The scVI latent provides a denoised representation for comparing chemistry-linked states across cell types. Gene-level recovery was evaluated separately with the exploratory decoder analysis.

Centered-response benchmark

Cell lines differ in baseline transcriptional state. A raw-space model can therefore score well by learning constant cell-line offsets. For the primary benchmark, source and target training means were subtracted before model fitting and evaluation. This centered-response formulation focuses performance on perturbation-linked variation and is the closest available proxy for response in the absence of matched vehicle controls in the canonical dataset.

Molecule-level transfer

For each library, cell line, and compound identifier, scVI latent coordinates were averaged across wells. Shared compounds between a source and target cell line were split into training and held-out test sets. Identity, global-shift where applicable, ridge transfer, PLS transfer, and k-nearest-neighbor transfer were compared. Performance was measured as row-wise cosine similarity on held-out compounds and averaged across repeated splits. The manuscript reports centered-response results as the primary molecule-level analysis and raw-profile results as supplementary diagnostics.

Grouped building-block transfer

Within each library and cell-pair direction, compounds were grouped by candidate building-block keys such as `bb0_id`, `bb1_id`, and multi-column combinations such as `bb0_id+bb1_id`. For each shared group, source and target centered latent signatures were averaged. A centered ridge transfer model was fit across repeated train-test splits, and performance was summarized as support-weighted cosine similarity across held-out groups. The main manuscript reports aggregation levels that remain chemically interpretable for the corresponding library design.

ZEL028-2 aggregation guardrail

ZEL028-2 grouped-transfer outputs were reviewed separately because single-building-block centroids in a four-position library collapse many distinct tuples. The guardrail table reports both high-scoring single-building-block rows and more specific BB-pair rows. Single-building-block rows are treated as broad diagnostics rather than chemistry-resolved program claims.

Cross-library projection

For a fixed source-target cell-line direction, a centered ridge transfer matrix was trained in one library and applied to another library. Because sparse libraries are singleton-heavy, adaptive minimum-well thresholds were used to retain feasible exploratory comparisons while preserving stricter filters in denser settings. Directional asymmetry was reported directly because teacher-library depth, target-library support, and chemistry-domain match are part of the result.

Exploratory gene-level decoder

For the four dense paired directions, a source-to-target latent transfer map was fit on shared training compounds. A separate target-cell ridge decoder was trained to map target-cell scVI centroids to log-CPM pseudobulk deltas over 800 variable genes selected from target training cells. Held-out predictions were scored by per-compound Pearson and Spearman correlation over gene deltas and by top-50 and top-100 absolute-delta gene overlap. This analysis was used for ranked gene-hypothesis generation.

Limitations

- Matched vehicle controls are unavailable in the canonical dataset, so centered response is a proxy for perturbational response rather than a fully DMSO-normalized effect.
- The primary molecule-level benchmark is limited to the dense paired libraries ZEL024 and ZEL031; ZEL028-2 becomes too sparse under strict per-molecule support filters.
- The strongest current transfer unit is the chemistry-resolved grouped program, with molecule-level transfer acting as supporting prioritization evidence.
- ZEL028-2 single-building-block results are intentionally treated as aggregation guardrails because they over-aggregate a four-position library.
- Cross-library projection depends on teacher-library depth, target-library support, and chemistry-domain match; the prospective version should include planned shared chemistry and balanced sampling.
- Gene-level decoding should be read as ranked hypothesis generation, not full transcriptome reconstruction.
- Current transfer models use averaged 32-dimensional scVI latent signatures and do not model dose, time, metabolism, tissue architecture, or organism-level context.

Data availability

Data tables, manuscript figures, and analysis inputs needed to reproduce this manuscript are organized under `paper4/` and `data/ZScreen_Canonical_Dataset/` in this repository. The public bundle replaces raw Z-Screen SMILES with `smiles_hash` identifiers and precomputed irreversible chemistry embeddings as described in the repository README. A persistent-archive deposition with an assigned DOI will accompany the corresponding preprint posting.

Code availability

Analysis and figure-generation scripts are in `paper4/scripts/`. Package-level dependencies are in the repository root `requirements.txt`. The package was tested with Python 3.14.3 on Windows.

References

1. Lopez R, Regier J, Cole MB, et al. Deep generative modeling for single-cell transcriptomics. Nat Methods. 2018;15:1053-1058. doi:10.1038/s41592-018-0229-2
2. Subramanian A, Narayan R, Corsello SM, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017;171(6):1437-1452.e17. doi:10.1016/j.cell.2017.10.049
3. Tahoe Therapeutics. Tahoe-100M: A 100-million single-cell perturbational atlas across 50 cancer cell lines. Dataset card and February 2025 release. https://huggingface.co/datasets/tahoe-bio/Tahoe-100M
4. Dapello J, Nassar M, Eksi R, et al. scGeneScope: A treatment-matched single-cell imaging and transcriptomics dataset and benchmark for treatment response

modeling. NeurIPS 2025 Datasets and Benchmarks. OpenReview. https://openreview.net/pdf/f7d541dae38bcf88a79789a4c6440aadfec123c7