LACK OF BROAD CROSS-SCREEN ALIGNMENT BETWEEN SC-ISLET TRANSPLANT RECOVERY AND BETA-CELL CYTOKINE-STRESS NECESSITY

OpenAI 5.6 Sol
GloGlo Research Team

Abstract. We tested whether relative guide recovery in a whole-genome CRISPR-activation SC-islet transplantation screen aligns with intact-gene necessity in an independent beta-cell cytokine-stress knockout screen. Across 17,286 joined genes, rank concordance was nearly zero ($\rho = 0.00660$, matched $P = 0.3798$). None of 3,132 biological-process modules passed the joint multiplicity, strongest-gene-removal, and stability gates; independent reconstruction preserved the result under stricter guide universes and fresh randomization. The public transplant counts also collapse twenty grafts into one kidney column, preventing animal-level consistency analysis. These data reveal no broad shared gene program between the two aggregate endpoints. Transplant recovery and cytokine-stress necessity should remain separate measured objectives.

1 Introduction

Genome-scale perturbation screens can measure distinct parts of the cell-therapy problem. A recent CRISPR-activation screen followed stem-cell-derived islets (SC-islets) through late differentiation, transplantation into immunodeficient mice, and recovery from kidney grafts ten days later [1]. An earlier knockout screen measured differential guide abundance in human EndoC-betaH1 cells exposed to a cytokine mixture relative to matched untreated cultures [2]. Both studies are relevant to type 1 diabetes cell replacement, but their perturbations, cellular systems, and aggregate outcomes differ.

A broad cross-screen alignment would be useful. It would indicate that genes or biological processes favored by activation during transplantation tend also to be required under inflammatory stress. Such an alignment could support shared mechanism models and common assay covariates. Its absence would imply that the two screens should not be treated as interchangeable measurements of a single latent "resilience" phenotype.

We therefore asked a restricted question: does the complete continuous transplant-recovery screen align with intact-gene necessity in the cytokine screen at either the gene or Gene Ontology biological-process level? The analysis plan fixed the score directions, exact joining, matched nulls, multiplicity control, strongest-gene removal, growth-confound sensitivity, and decision thresholds before the transplant source values or ranked results were inspected. A separate implementation then audited the source units, identifiers, randomization, and result.

Date: August 30, 2026.
Key words and phrases. stem-cell-derived islets, CRISPR activation, transplantation, cytokine stress, gene-set analysis, matched randomization.
OpenAI 5.6 Sol developed the analysis and performed the accompanying verification. The GloGlo research team posed the question, established the research direction, and guided the investigation.

2 Data and methods

2.1 SC-islet transplantation screen

The source was the complete workbook accompanying Maestas et al. [1]. CRISPR activation was introduced before SC-islet differentiation, and the screen compared day-0 abundance with abundance after ten days in kidney-capsule grafts in NSG mice. We oriented the transplant score so that positive values indicated greater relative post-transplant recovery. For targeting annotation row i, the guide score was

$$s_i = \log_2 \left(\frac{\mathrm{CPM}_{i,\mathrm{kidney}} + 0.5}{\mathrm{CPM}_{i,\mathrm{D0}} + 0.5} \right),$$

and target-level scores were the median of eligible guide scores without selection on the reported result.

The workbook contains 106,400 annotation rows: 102,645 targeting rows and 3,755 nontargeting controls. Independent sequence-level reconstruction found 103,090 unique guide sequences in all, including 99,335 unique targeting sequences. Of these, 3,074 sequences were assigned to more than one target symbol, generating 6,328 duplicated annotations and touching 2,166 symbols. The primary reconstruction retained the deposited annotation structure. Two audit sensitivities excluded multi-target sequences and, more strictly, excluded targets with multiple promoter units.

The publication reports twenty transplanted mice, whereas the public workbook contains one day-0 count column and one kidney count column. We consequently analyzed one aggregate comparison. Grafts were not treated as independent replicates.

2.2 Cytokine-stress intact-gene necessity

The second source was the complete deposited gene-level result from the EndoC-betaH1 CRISPR-knockout cytokine screen, GEO GSE205853 [2]. Its positive knockout effect denotes enrichment after cytokine exposure. To align directions across perturbations, we defined intact-gene necessity as

$$n_g = \mathrm{clip}(-z_g, -4, 4) \sqrt{\frac{\min(k_g, 6)}{6}},$$

where z_g is the deposited gene statistic and k_g the number of guides. Thus a positive cross-screen association would mean that activation favored relative transplant recovery while knockout was depleted under cytokine stress. It would be cross-context coherence, not experimental replication.

Exact symbol joining yielded 18,407 rows. Requiring a unique historical symbol-to-Ensembl mapping produced the 17,286-gene primary universe; 15,497 of these genes had an eligible frozen biological-process annotation. The initial machine registry omitted the historical symbol-map file even though the analysis used it. The independent audit pinned its bytes, reproduced the 17,286-gene join exactly, and records this provenance amendment.

2.3 Gene-level tests

The primary statistic was Spearman correlation between continuous transplant recovery and intact-gene necessity. Both measures were also converted to average-tie ranks and inverse-normal scores for a Pearson sensitivity. A two-sided 10,000-randomization test shuffled transplant scores within joint strata defined by annotation degree, guide support, common-essential status, and cell-cycle or proliferation status. A stratified 10,000-resample bootstrap supplied the interval for the primary correlation.

We additionally tested overlap between the top 5% and top 10% of each oriented score, using stratum-preserving randomization and Benjamini–Hochberg correction across correlation and overlap tests. A leave-one-out sensitivity removed the single gene with the largest absolute product of the two rank-normal scores, selected without reference to its identity.

2.4 Biological-process tests

Modules were direct human Gene Ontology biological-process memberships [3] plus Reactome-assigned GO memberships [4], frozen to September 1, 2021. Eligible modules contained 8–200 observed genes. Identical member sets were collapsed before testing. For module M, the conjunction statistic was

$$T_M = \min \left\{ |M|^{-1} \sum_{g \in M} z_g^{(T)}, |M|^{-1} \sum_{g \in M} z_g^{(C)} \right\},$$

where $z^{(T)}$ and $z^{(C)}$ are the transplant and cytokine-necessity rank-normal scores. Ten thousand matched random gene sets preserved observed counts within the joint strata. Full-set and strongest-contributing-gene- removed empirical P values were corrected separately across all modules.

A module required positive means in both sources, full and leave-one-out $q \leq 0.10$, at least 90% both-positive stability, and no dominant single-gene contribution. Stability used 2,000 draws of 80% of member genes without replacement. We refer to this operation as *subsampling*; the original analysis code called it a bootstrap. Independent audit also repeated the stability calculation by conventional sampling with replacement. Because no module passed both multiplicity and leave-one-out gates, the choice between these stability procedures could not change the result.

The growth-deconfounded sensitivity removed common-essential, cell-cycle, and cell-population-proliferation genes before repeating the gene and module analyses.

2.5 Independent reconstruction

The audit used a separate implementation, direct reconstruction from the pinned source workbook, fresh random streams, and three source universes: the reported annotation structure, an ambiguity-clean universe that removed every multi-target sequence, and a strict universe that additionally retained only single-promoter targets. It reproduced the deterministic outputs and ran fresh 10,000-randomization analyses for all six universe-by-confound combinations. A 50,000-randomization sensitivity assessed the Monte Carlo resolution of the module family.

3 Results

3.1 The public transplant object is a high-coverage aggregate

The deposited workbook contained 18,913 raw target symbols and 18,725 rows in the author's continuous RRA table. The author table was therefore treated as an orientation sensitivity rather than the primary universe. All raw target symbols had at least five annotation rows. After correcting the retention denominator to targeting rows detected at day 0, 102,599 of 102,604 rows were also detected after transplantation (0.999951). Sequence deduplication reduced the apparent D0 and kidney totals from 66,916,017 and 61,558,710 row-summed counts to 62,720,358 and 57,728,324 unique-sequence counts, respectively.

These diagnostics show extensive aggregate representation. They do not measure consistency across the twenty grafts, because no animal-indexed screen counts are public. Physical pooling before sequencing cannot be distinguished from reporting a collapsed aggregate.

3.2 Gene-level alignment is nearly zero

Across 17,286 exact-joined genes, the primary Spearman correlation was 0.00660 (matched $P = 0.3798$; Table 1). The inverse-normal Pearson correlation was 0.00698, and the stratified bootstrap interval for Spearman correlation was $[-0.00831, 0.02139]$. Removing the strongest product gene changed the correlation only to 0.00643 ($P = 0.3939$).

Table 1: Gene-level cross-screen alignment. Audit P values use fresh 10,000-randomization streams.

Analysis	Genes	Spearman ρ	Matched P
Frozen primary	17,286	0.00660	0.3798
Frozen growth-deconfounded	15,558	0.00472	0.5509
Audit ambiguity-clean	16,800	0.00815	0.2885
Audit ambiguity-clean, deconfounded	15,132	0.00486	0.5500
Audit single-promoter	15,339	0.00772	0.3437
Audit single-promoter, deconfounded	13,829	0.00572	0.4930

The top-tail results agreed with the continuous analysis. The top-5% sets overlapped at 45 of 865 genes (Jaccard 0.0267; matched family $q = 0.3835$). The top-10% sets overlapped at 179 genes (Jaccard 0.0546; matched family $q = 0.3835$). Removing growth-associated genes did not reveal an overlap or correlation. The ambiguity-clean and single-promoter audit universes likewise produced correlations between 0.00486 and 0.00815, all with intervals spanning zero.

3.3 No biological-process module passes the joint gates

The primary universe generated 3,132 unique biological-process modules; the growth-deconfounded universe generated 2,547. Zero modules passed the full analysis, and zero passed after strongest-gene removal. The fixed-seed 10,000-null run had minimum full and leave-one-out q values of 0.1566 and 0.7829 in the primary analysis, and 0.1273 and 1.0000 after growth deconfounding (Table 2).

Table 2: Module-family results. "Full" and "LOO" are the minimum Benjamini–Hochberg q values before and after strongest-gene removal.

Analysis	Modules	Min full q	Min LOO q	Joint passes
Frozen primary, 10,000 nulls	3,132	0.1566	0.7829	0
Frozen deconfounded, 10,000 nulls	2,547	0.1273	1.0000	0
Audit primary, fresh 10,000 nulls	3,132	0.2088	0.6263	0
Audit deconfounded, fresh 10,000 nulls	2,547	0.2547	0.2547	0
Audit primary, 50,000 nulls	3,132	0.1879	0.5637	0
Audit deconfounded, 50,000 nulls	2,547	0.0509	0.2038	0

At 50,000 randomizations, one growth-deconfounded module crossed the full-set $q \leq 0.10$ threshold, but failed the independently required strongest-gene-removed gate ($q = 0.2038$). No other leave-one-out module passed. This higher-resolution sensitivity therefore left the joint result

unchanged. Fresh ambiguity-clean and single-promoter analyses tested between 2,258 and 3,075 modules; none passed both full and leave-one-out multiplicity gates.

Subsampling and conventional bootstrap stability counts differed substantially: 251 versus 30 primary modules, and 197 versus 21 deconfounded modules, reached the 0.90 stability threshold. This difference exposes why the resampling unit must be stated precisely. It had no bearing on the terminal result because the multiplicity and leave-one-out gates had already failed.

4 Discussion

Two complete, continuously scored genome-scale screens showed no broad cross-screen alignment. The gene-level correlation was approximately zero, top-tail overlap was consistent with matched nulls, and no biological-process module passed the joint criteria. The conclusion survived removal of growth and essentiality genes, removal of ambiguous guide sequences, restriction to single-promoter targets, strongest-gene deletion, and independent null generation.

This result draws a boundary around two aggregate endpoints. Relative guide recovery after SC-islet differentiation and transplantation can reflect activation, developmental state, endocrine composition, proliferation, engraftment, vascularization, sampling, or cell loss. The cytokine screen uses a different beta-cell system, knockout rather than activation, and a treated- versus-untreated abundance contrast. A null cross-screen correlation is therefore compatible with context-specific biology in either experiment. It shows that the two measurements do not presently reveal one broad shared gene-ranking or biological-process axis.

The practical implication is architectural. Transplant recovery and cytokine-stress necessity should be modeled as separate objectives rather than collapsed into a single resilience label. Dynamic insulin secretion, beta-cell identity, absolute cell fate, and immune evasion are additional non-substitutable outcomes. Optimizing or validating a model on one aggregate screen does not establish performance on the others.

The source audit also changes the minimum data object needed for the next analysis. Per-animal guide-count matrices, sample provenance, and an explicit pooling history would allow animal-level consistency and graft-bottleneck estimation. Direct measurements of activation, absolute graft-cell number, early death, endocrine composition, and dynamic secretion would separate the mechanisms compressed into relative recovery. Immune-evasion questions require an immune-competent or defined immune-effector context rather than the NSG transplant screen used here.

5 Limitations

The transplant screen is represented publicly by one day-0 and one kidney count column. Consequently, neither between-graft variance nor animal-level effect consistency is identifiable. The deposited annotation table contains multi-target guide sequences and multiple promoter units; although stricter reconstructions preserve the result, they reduce the eligible universe. The historical symbol map used in the primary exact join was pinned only during the independent audit rather than in the initial machine registry.

The analysis compares marginal gene scores across two different perturbation directions, cell systems, and experimental environments. A nonlinear, conditional, or small gene-specific relationship could exist without producing genome-wide rank concordance. GO module tests depend on the frozen annotation version and on mean rank-normal scores. Finally, both source outcomes are relative abundance measurements; neither alone decomposes death, division, state transitions, and sampling.

6 Conclusion

The complete public data provide no evidence for a broad shared gene program between SC-islet transplant recovery after CRISPR activation and intact-gene necessity under beta-cell cytokine stress. This is a result about alignment between two aggregate endpoints, not about the value of either screen in its own context. Future computational and experimental programs should preserve transplant fitness, inflammatory stress response, beta-cell function, and immune interaction as distinct measured objectives.

Data and code availability

The transplant source workbook accompanies Maestas et al. [1]; the cytokine-screen result is available under GEO accession GSE205853 [2]. The preregistration, reconstruction, independent audit, exact input hashes, machine-readable outputs, and replay instructions are contained in the GloGlo research repository under `analyses/scislet_crispra_transplant_module_map_v1` and `analyses/scislet_crispra_transplant_module_map_independent_audit`.

References

[1] M. M. Maestas, K. Bradley, M. Shunkarova, et al., "Whole-genome CRISPR screening identifies genetic modifiers of stem cell-derived islet transplantation," *Stem Cells Translational Medicine*, vol. 15, no. 4, szag012, 2026. doi:10.1093/stcltm/szag012.

[2] P. Benaglio, K. A. D'Antonio-Chronowska, W. Ma, et al., "Genetic and epigenetic regulation of human beta-cell responses to inflammatory stress," *Cell Genomics*, vol. 2, no. 12, 100214, 2022. doi:10.1016/j.xgen.2022.100214.

[3] The Gene Ontology Consortium, "The Gene Ontology resource: enriching a GOld mine," *Nucleic Acids Research*, vol. 49, D325–D334, 2021. doi:10.1093/nar/gkaa1113.

[4] B. Gillespie, B. Jassal, R. Stephan, et al., "The Reactome pathway knowledgebase 2022," *Nucleic Acids Research*, vol. 50, D687–D692, 2022. doi:10.1093/nar/gkab1028.