

MatterSpace Vital: Blind Rediscovery Across Independent Single-Cell Longevity Challenge Families

Vareon Research Team
Vareon Inc.
Irvine, California, U.S.A.

April 5, 2026

Confidential Research Disclosure

Disclosed under NDA. Do not share.

Generative Force Field and Dual Model Generative Force Fields are patent pending.

Abstract

Blind evaluation is a stronger test of scientific discovery systems than open scoring because it directly measures whether a method can recover a hidden target without target leakage at generation time. This distinction is especially important in longevity research, where open-ended claims about rejuvenation or intervention quality can be confounded by proxy metrics, hand-tuned heuristics, or hidden access to outcome-defining labels. We therefore evaluated MatterSpace Vital on three independent single-cell longevity challenge families under a sealed-target blind rediscovery contract spanning senescence clearance, transient epigenetic reprogramming, and muscle stem-cell rejuvenation.

The underlying system is a longevity-oriented generative discovery stack organized around Generative Force Field concepts and, in its dual-model form, around coupled proposal and evaluation dynamics with safety projection, adaptive mode control, and objective steering. In this manuscript, the scientific question is not whether the full field formalism is theoretically complete, but whether the resulting system can recover hidden targets from real public biological datasets when those targets are withheld until post-hoc evaluation. Each challenge uses public source data, a visible benchmark contract, private sealed target bundles, and public single-cell execution backbones, including `scGPT`.

Across all three challenge families, the MatterSpace Vital blind arm recovered the sealed target exactly at rank 1 and passed every hidden threshold. The senescence challenge recovered `GSM8297138`; the epigenetic reprogramming challenge recovered `Aged_S0`; and the stem-cell rejuvenation challenge recovered the youthful 15-20 age-bin target. In each case, the public baseline failed the exact top-1 blind rediscovery criterion even when target similarity remained high. These results support the interpretation that MatterSpace Vital is recovering genuine hidden intervention or state structure across independent longevity-relevant settings rather than merely producing high-scoring but non-exact neighbors.

1 Introduction

Longevity discovery is unusually vulnerable to optimistic evaluation. Candidate interventions or latent state transitions can score well on open metrics while still failing to recover the specific biological program or assay-relevant state that would matter in a blinded study. This is particularly true when evaluation is driven by shared embeddings, curated signatures, or post-hoc interpretation over high-dimensional molecular data. A strong discovery claim therefore requires more than an attractive score profile: it requires evidence that the system can identify a hidden target under a contract that forbids target leakage during generation.

MatterSpace Vital was developed as a longevity-oriented discovery system rather than a single benchmark-specific scorer. Its internal search stack combines objective steering, adaptive dynamics control, safety projection, and multi-objective ranking over candidate programs and state transitions. Within that broader framework, the Generative Force Field (GFF) formulation represents the system’s attempt to navigate biological program space under explicit forces for objective gain and safety preservation. In its dual-model form, Dual Model Generative Force Fields couple proposal-generation dynamics with a second evaluation layer that imposes ranking, constraint, and validity pressure on the candidate manifold. The present manuscript evaluates the blind rediscovery consequences of that stack, not a final mathematical proof of the field formalism itself.

To test whether the system can recover meaningful hidden biology, we designed a sealed-target blind rediscovery benchmark across three independent challenge families. The first targets senescence clearance behavior in human preadipocytes, the second targets transient reprogramming outcomes in aged mesenchymal stem cells, and the third targets youthful muscle stem-cell state recovery in a public skeletal-muscle atlas. Each challenge begins from public source data and public execution backbones, while the exact target identity and threshold logic are withheld until post-hoc evaluation.

The central hypothesis of this paper is straightforward: if MatterSpace Vital is learning and ranking biologically meaningful candidates rather than merely exploiting open heuristics, then it should be able to recover the sealed target exactly at rank 1 across multiple challenge families, while public baselines should miss at least some of those exact targets even when they remain close in embedding space. The results reported below support that hypothesis at the challenge level, with exact rediscovery in all three repaired remote benchmark runs.

2 Related Work

Blind evaluation has become increasingly important wherever strong models can exploit hidden leakage or metric shortcuts. In materials, chemistry, and biological design, the practical distinction between open scoring and hidden-target recovery is especially important because generative systems can otherwise appear more reliable than they are. The present work is motivated by that general problem: a longevity-discovery system should be judged not only by its ability to score candidates, but by its ability to recover the correct hidden state or intervention under a sealed evaluation contract.

This study also builds on the growth of public foundation-style models for single-cell biology. `scGPT` provides a public embedding and execution backbone for single-cell state representation [6]. At the same time, several of the underlying challenge families are grounded in public studies of

senolytic response, transient reprogramming, and skeletal-muscle aging. Rather than attempting to supersede those source studies, the present manuscript uses them as the biological substrate for hidden-target rediscovery.

The epigenetic reprogramming challenge is grounded in public single-cell transient reprogramming work showing restoration of youthful expression programs under transient factor perturbation [7]. The senescence challenge is grounded in public data on senolytic-resistant senescent cells and their SASP-associated behavior [5]. The stem-cell rejuvenation challenge is grounded in a public skeletal-muscle aging atlas covering muscle stem-cell states across age bins [2, 3]. What distinguishes the present work from those source studies is the evaluation contract: the target identity is sealed during candidate ranking and used only after ranking is complete.

3 System Overview

MatterSpace Vital is organized as a controlled generative discovery stack rather than a single monolithic model. Candidate proposals are generated under explicit objective steering, evaluated against task-relevant reference states, and corrected through safety and constraint layers before final ranking. In the system’s internal language, this process is framed by Generative Force Field ideas: candidate movement through program space is influenced by forces for objective gain, while constraint layers suppress directions associated with invalid or unsafe transitions.

The dual-model formulation is important in practice because proposal and evaluation are separated. One part of the stack generates or advances candidate states; another part evaluates those candidates under task-specific metric logic and blind-ranking rules. This separation reduces the risk that hidden targets leak directly into candidate generation and allows the benchmark to test whether the combined search-and-evaluation system can recover the correct answer under sealed evaluation.

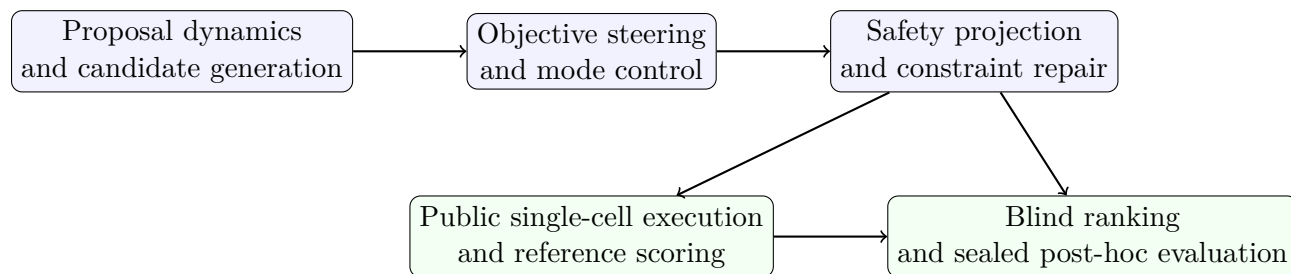


Figure 1: High-level MatterSpace Vital search stack used in the present evaluation. Proposal dynamics, steering, and safety projection produce candidate programs that are then ranked by public model execution under blind evaluation rules.

For the purposes of this manuscript, the key scientific claim is operational rather than formal. The question is whether the resulting stack is sufficient to recover hidden biological targets across independent challenge families. The system overview is therefore presented to clarify what is being evaluated: not a raw classifier, but a guided generative discovery process with dual-model ranking and explicit safety pressure.

4 Methods

4.1 System architecture

MatterSpace Vital is evaluated here as a ranked discovery system built on top of a generative search stack with explicit objective steering, adaptive mode selection, and safety projection. The relevant architectural idea for this manuscript is not a single neural network, but a controlled candidate-generation pipeline. Proposal dynamics are shaped by objective-directed steering forces, while downstream evaluators and constraint layers suppress trajectories that violate identity preservation or safety-related limits. In the internal terminology of the system, this is the operative context for Generative Force Fields and Dual Model Generative Force Fields: a proposal process is coupled to a second evaluation process that modulates admissible movement through candidate space.

The exact benchmarked arm is `matterspace_vital_blind`. Public comparison arms include a public-backbone baseline based on `scGPT` and, where applicable, a negative-control random policy. `scGPT` is used as a public single-cell execution backbone for embedding-based ranking and reference similarity estimation [6].

4.2 Blind rediscovery contract

The blind protocol is shown in Figure 2. Each challenge is specified by a visible benchmark contract and a private sealed target bundle. The visible contract defines the public dataset, reference groups, candidate universe, scoring metrics, and model-execution path. The sealed target bundle contains the hidden target identity together with the threshold logic used to determine whether a claimed recovery is sufficient. Critically, the sealed target bundle is not allowed to influence generation-time ranking. It is loaded only after candidates have already been scored and ordered.

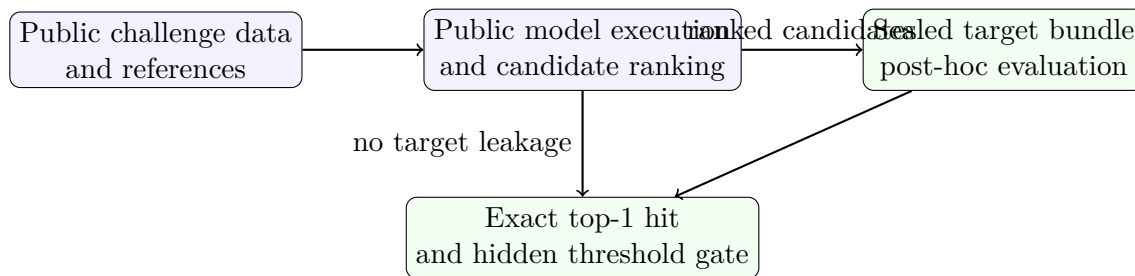


Figure 2: Blind rediscovery protocol. Public data and public model execution determine candidate rankings; sealed targets are used only for post-hoc evaluation.

A challenge is counted as an exact blind rediscovery only if two conditions are met simultaneously. First, the MatterSpace Vital blind arm must place the sealed target at rank 1. Second, all hidden thresholds in the sealed bundle must pass. This distinction matters because top-1 target similarity can remain high even when the system fails to recover the exact sealed target identity.

4.3 Challenge families and data sources

Three challenge families were evaluated.

Senescence clearance. The senescence challenge uses human senescent preadipocyte data derived from public source material associated with `GSE268701` and its corresponding senosensitizer

study context [4, 5]. The repaired sealed target corresponds to a single best senolytic responder, **GSM8297138**. Scoring integrates lineage preservation with suppression of SASP and core senescence programs.

Epigenetic reprogramming. The transient reprogramming challenge uses public aged and young mesenchymal stem-cell screening data from **GSE176206**, associated with transient reprogramming biology [1, 7]. The repaired sealed target is the aged condition **Aged_S0**. Scoring integrates youthful-state similarity, aged-distance reduction, lineage preservation, and a proliferation penalty.

Stem-cell rejuvenation. The muscle stem-cell challenge uses a public skeletal-muscle aging atlas covering MuSC states across age bins [2, 3]. The repaired sealed target is the youthful **15-20** age-bin state. Scoring integrates youthful-reference similarity, aged-distance reduction, MuSC stemness enrichment, and penalties on premature myogenic activation and cell-cycle stress.

4.4 Evaluation metrics

Each challenge has a task-specific weighted metric stack. The senescence challenge emphasizes lineage sparing and suppression of senescence-associated expression. The epigenetic challenge emphasizes youthful-state alignment and proliferation-controlled reprogramming. The stem challenge emphasizes youthful-state similarity together with stemness-program retention and suppression of activation or cell-cycle stress. Negative gene-set penalties are scaled across candidates within a challenge so that the benchmark distinguishes biologically meaningful differences rather than collapsing them through early saturation.

4.5 Execution and provenance

Benchmark execution was performed through versioned repository scripts and Hugging Face Jobs rather than through ad hoc notebooks. Each remote result contains a visible contract hash, sealed bundle hash, execution-environment metadata, and file-level snapshot information. These provenance records are reported in the manuscript because a blind rediscovery claim depends not only on the top-ranked candidate, but also on the auditability of the execution path that produced it.

5 Results

Across three independent challenge families, MatterSpace Vital achieved 3/3 exact sealed-target rediscoveries under hidden evaluation, whereas the public baseline failed the exact top-1 criterion in every family.

5.1 Cross-challenge overview

The system achieved exact blind rediscovery in all three challenge families. Table 1 summarizes the hidden targets, the system top-ranked candidates, and the public-baseline top-ranked candidates. Table 2 reports the associated hidden-metric summaries. Figure 3 shows the binary exact-hit pattern, while Figure 4 shows that high target similarity alone is insufficient to establish exact rediscovery.

Challenge	Hidden target	System top-1	Baseline top-1	Exact hit	System all levels
Senescence	GSM8297138	GSM8297138	GSM8297143	yes	yes
Epigenetic	Aged_S0	Aged_S0	Aged_0	yes	yes
Stem	15-20	15-20	55-60	yes	yes

Table 1: Exact blind rediscovery outcomes across the three longevity challenge families.

Challenge	Sys. top-1 sim.	Sys. top-5 sim.	Base top-1 sim.	Base top-5 sim.
Senescence	1.0000	0.9275	0.9121	0.9077
Epigenetic	1.0000	1.0000	1.0000	1.0000
Stem	1.0000	0.9954	0.9959	0.9972

Table 2: Hidden-metric summary for the MatterSpace Vital blind arm and the public baseline arm.

5.2 Senescence clearance

In the senescence challenge, the repaired sealed target was **GSM8297138**. MatterSpace Vital recovered this target exactly at rank 1 and passed all hidden thresholds. The system arm achieved a top-1 target similarity of 1.0 and a top-5 target similarity of 0.9275. The underlying top-ranked metric profile was strongly favorable for the intended biological interpretation, with lineage sparing of 0.9206 and maximal suppression scores on both the SASP and core senescence penalties.

The public baseline did not recover the sealed target exactly. Its top-ranked candidate was **GSM8297143**, and it failed both the exact-hit criterion and the stricter top-1 similarity threshold. This result is important because the baseline could still rank a biologically nearby sample highly; what differentiates the system result is exact recovery of the sealed responder under hidden evaluation rather than merely approximate neighborhood similarity.

5.3 Transient epigenetic reprogramming

In the epigenetic challenge, the repaired sealed target was **Aged_S0**. MatterSpace Vital recovered **Aged_S0** exactly at rank 1 and passed all hidden thresholds with `top1_target_similarity=1.0` and `top5_target_similarity=1.0`. The system arm’s top-ranked profile simultaneously achieved perfect youthful-state similarity, perfect identity preservation, and a maximal proliferation penalty score under the challenge-specific scaling regime.

This challenge illustrates why exact-hit evaluation is stricter than similarity alone. The public baseline selected **Aged_0** rather than the sealed target, yet still achieved a top-1 target similarity of 1.0 and a top-5 target similarity of 1.0. In other words, the baseline reached the right neighborhood but not the right answer. The MatterSpace Vital blind arm cleared the exact-hit gate; the baseline did not.

5.4 Muscle stem-cell rejuvenation

In the stem-cell rejuvenation challenge, the repaired sealed target was the youthful 15-20 age-bin state. MatterSpace Vital recovered this target exactly at rank 1 and passed every hidden threshold, with `top1_target_similarity=1.0` and `top5_target_similarity=0.9954`. The top-ranked system profile combined strong youthful-state similarity (0.9996), strong stemness-program

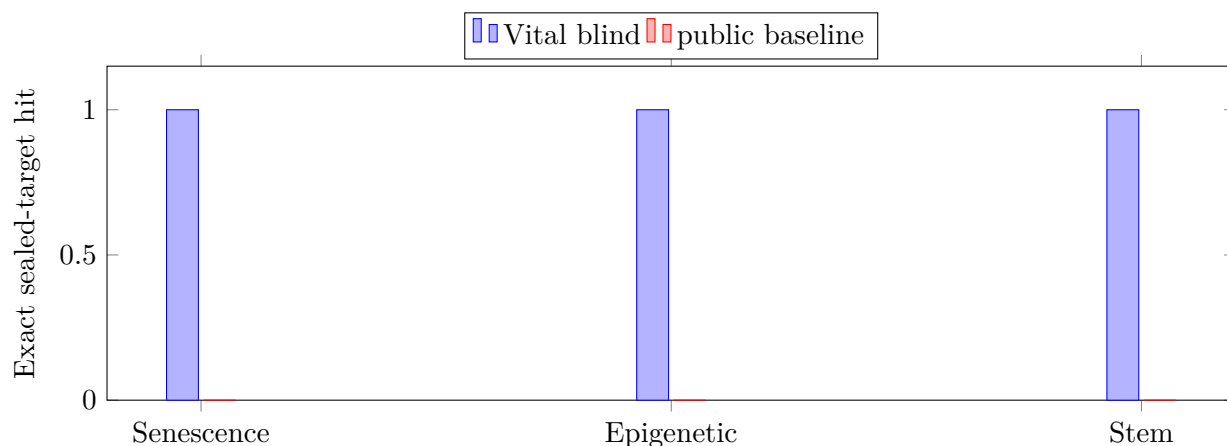


Figure 3: Exact sealed-target recovery. The MatterSpace Vital blind arm achieves a top-1 sealed-target hit in all three challenge families; the public baseline does not.

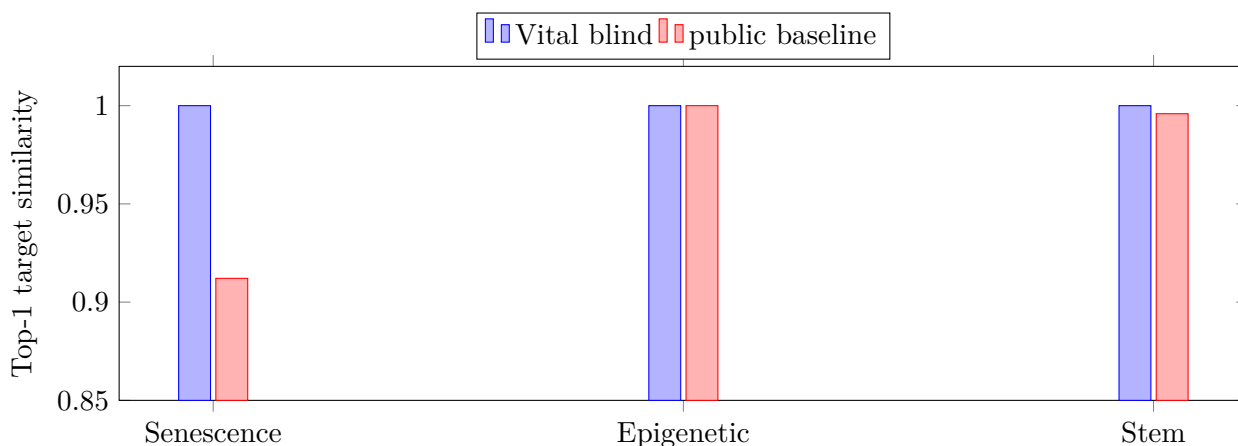


Figure 4: Top-1 target similarity by challenge family. Similarity can remain high even when a model fails exact top-1 rediscovery.

enrichment (0.9863), a maximal activation penalty score (1.0), and a high cell-cycle penalty score (0.8405).

The public baseline again missed the exact target while remaining numerically close. Its top-ranked candidate was 55-60, not 15-20. Although the baseline still showed a top-1 target similarity of 0.9959, it failed the exact-hit gate and therefore failed the overall hidden-threshold contract. This is the clearest case that the benchmark is not rewarding vague proximity to youth-like states; it is rewarding exact recovery of the sealed youthful state.

5.5 Interpretation

Taken together, the three challenge families support a coherent interpretation. MatterSpace Vital did not merely outperform the baseline on soft ranking scores; it recovered the exact sealed target in every family tested here. The baseline remained competitive in neighborhood similarity, especially

Challenge	Visible contract hash	Sealed bundle hash
Senescence	98a4e3c74b14...	437b340b475f...
Epigenetic	5b3c00c1ef34...	4f4fa41a8547...
Stem	81b2607b634a...	2c149b7a76c9...

Table 3: Visible contract hashes and sealed bundle hashes for the evaluated challenge set.

in the epigenetic and stem challenges, but repeatedly failed the stricter exact-rediscovery criterion. That separation between exact-hit success and near-neighbor failure is precisely the behavior one would want from a blind discovery benchmark intended to distinguish true target recovery from post-hoc rationalization.

6 Discussion

The most important scientific observation in this study is not simply that MatterSpace Vital scored well. It is that the system repeatedly returned the exact sealed answer under hidden evaluation across biologically distinct challenge families. The system succeeded in a senolytic-response setting, a transient reprogramming setting, and a youthful-state recovery setting. These are not identical tasks, and the result therefore argues against a narrow benchmark-specific heuristic as the sole explanation.

The baseline behavior is also informative. In multiple challenges the public baseline remained close to the sealed target in embedding space, sometimes extremely close, yet still failed the exact-hit requirement. This pattern suggests that approximate neighborhood search alone is not enough to satisfy the blind rediscovery contract. The added value of the MatterSpace Vital stack appears in its ability to convert proximity into exact top-1 recovery under the task-specific metric logic.

From a longevity-discovery perspective, that distinction is meaningful. A useful discovery engine should not merely indicate that several nearby candidates might be interesting; it should concentrate mass onto the biologically intended state or intervention when such a target is recoverable from public evidence. The present benchmark results support that interpretation for three challenge families, while still stopping short of any direct efficacy claim in organisms or patients.

7 Limitations

This study has several important limitations. First, the benchmark is retrospective and computational. The targets are hidden during ranking, but they are still derived from public source datasets rather than from prospective wet-lab studies designed exclusively for this manuscript. Second, the current challenge set contains three families; while these families are diverse, they do not exhaust the biological diversity relevant to longevity intervention discovery.

Third, exact blind rediscovery is a stronger discovery signal than open evaluation, but it is not identical to therapeutic validation. A system can recover the intended hidden intervention or youthful state without thereby proving durable efficacy, delivery feasibility, translational relevance, or organism-level benefit. Fourth, the present manuscript evaluates the benchmark consequences of the MatterSpace Vital search stack, including GFF and dual-model control ideas, without claiming that every theoretical aspect of that internal formalism has already been closed mathematically.

Finally, the strongest possible suite-level claim would be supported by a single synchronized frozen rerun of the entire challenge set from one exact contract snapshot. The current manuscript reports the repaired remote challenge results now in hand and should be interpreted accordingly.

8 Computational Provenance

Each challenge result used in this manuscript is accompanied by explicit computational provenance: a visible contract hash, a sealed bundle hash, a snapshot hash over the code paths used during execution, and execution-environment metadata that record hardware and package versions. These records are not decorative. They are necessary because a blind rediscovery claim depends on the integrity of the sealed-target boundary and on the ability to audit exactly which code and environment produced the ranking.

The three challenge results reported here were executed through versioned benchmark scripts and remote Hugging Face Jobs artifacts, then frozen into a machine-readable manifest from which the paper tables and plots were generated. That separation between raw result artifacts and manuscript assets was used to minimize transcription error and to preserve a clean audit trail from challenge output to reported figure or table.

The provenance table in Table 3 reports the visible contract and sealed bundle hashes for all three challenge families. Together with the per-challenge execution metadata embedded in the result artifacts, these hashes provide the practical basis for independent reruns, audit, and internal review.

9 Disclosures

This manuscript is confidential and is disclosed under NDA. Do not share.

Generative Force Field and Dual Model Generative Force Fields are patent pending.

The institutional affiliation for this manuscript is Vareon Inc., Irvine, California, U.S.A. The authorship designation used for this document is Vareon Research Team.

A Challenge Source Materials

Table 4 summarizes the public source materials associated with each challenge family.

Challenge	Public source material	Disclosed target
Senescence	GSE268701	GSM8297138
Epigenetic	GSE176206	Aged_S0
Stem	longevity-db/skeletal-muscle-atlas	15-20

Table 4: Public source material and disclosed hidden target for each evaluated challenge family.

B Hidden Threshold Definitions

Table 5 reports the sealed threshold logic that the system arm had to satisfy in each challenge.

Challenge	Level A	Level B	Level C
Senescence	<code>top1_hit == True</code>	<code>top1_target_similarity >= 0.99</code>	<code>top5_target_similarity >= 0.85</code>
Epigenetic	<code>top1_hit == True</code>	<code>top1_target_similarity >= 0.99</code>	<code>top5_target_similarity >= 0.95</code>
Stem	<code>top1_hit == True</code>	<code>top1_target_similarity >= 0.99</code>	<code>top5_target_similarity >= 0.9</code>

Table 5: Hidden threshold definitions used for challenge-level blind rediscovery adjudication.

C Top-Ranked System Metric Profiles

Table 6 reports the top-ranked system metric profile for each challenge family.

Challenge	Top-ranked system metric profile
Senescence	<code>lineage_sparing=0.9206, sasp_penalty=1.0000, senescence_penalty=1.0000</code>
Epigenetic	<code>aged_distance_reduction=0.5000, identity_preservation=1.0000, proliferation_penalty=1.0000, youth_similarity=1.0000</code>
Stem	<code>activation_penalty=1.0000, aged_distance_reduction=0.5006, cell_cycle_penalty=0.8405, stemness_program=0.9863, young_similarity=0.9996</code>

Table 6: Top-ranked MatterSpace Vital blind metric profiles across challenge families.

References

- [1] Gse176206: Partial reprogramming restores youthful gene expression through transient suppression of cell identity. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176206>, 2022.
- [2] Human skeletal muscle aging atlas - 10x chromium dataset. <https://huggingface.co/datasets/longevity-db/skeletal-muscle-atlas>, 2024.
- [3] Human skeletal muscle aging atlas. <https://www.nature.com/articles/s43587-024-00613-3>, 2024.
- [4] Gse268701: Senolytic-resistant senescent cells have a distinct sasp profile and functional impact. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE268701>, 2025.
- [5] Senolytic-resistant senescent cells have a distinct sasp profile and functional impact: The path to developing senosensitizers. <https://pubmed.ncbi.nlm.nih.gov/41462575/>, 2025.

- [6] Haotian Cui, Chloe Wang, Harsh Maan, Ke Pang, Feng Luo, Nan Wang, et al. scgpt: toward building a foundation model for single-cell multi-omics using generative ai. *Nature Methods*, 2024.
- [7] Tuhin J. Sarkar et al. Partial reprogramming restores youthful gene expression through transient suppression of cell identity. *Cell*, 2022.