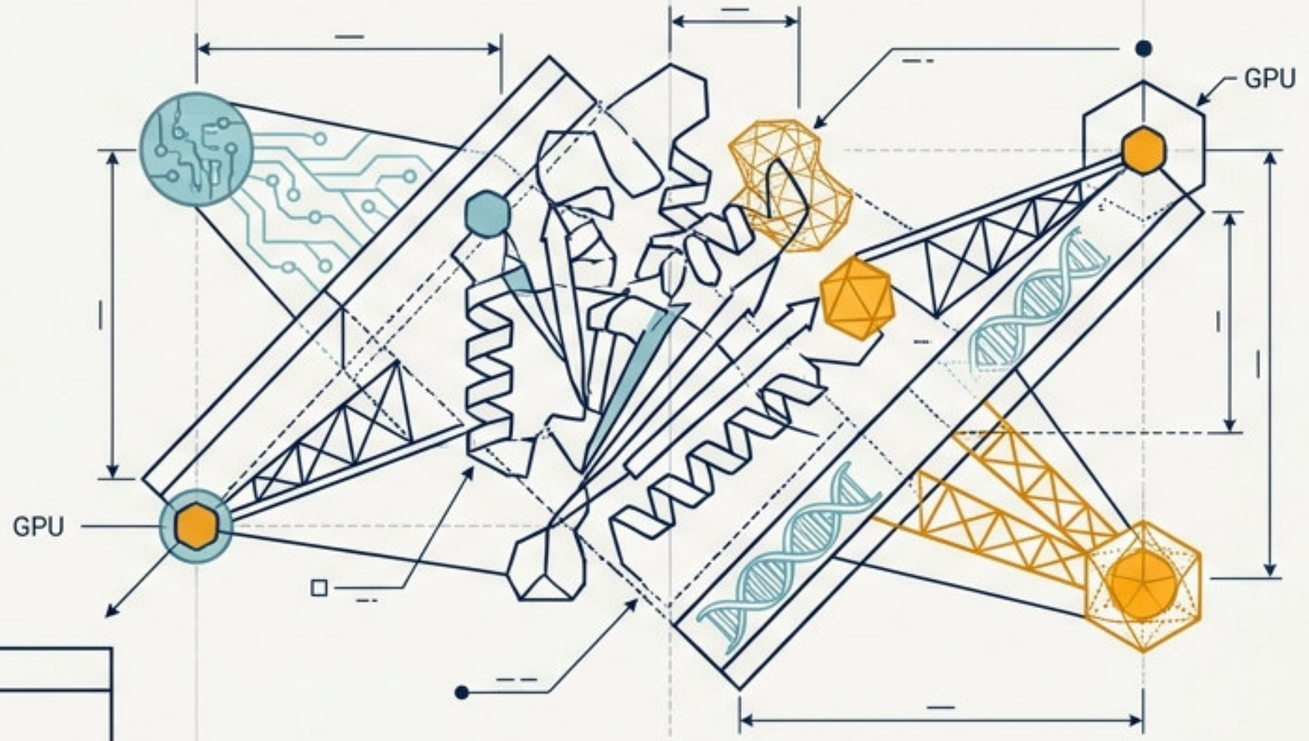
# The Spacefarer Phenome

An AI-orchestrated, repurposing-first computational drug-discovery pipeline for astronaut resilience.

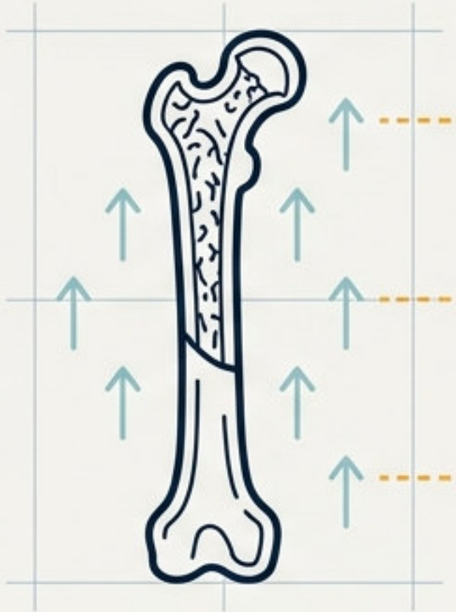



```
SYS_RUN: 2026-07-12 |
BUILD: GRCh38 |
STATUS: FROZEN
```

NotebookLM

# Microgravity as a Catalyst for Biological Ageing

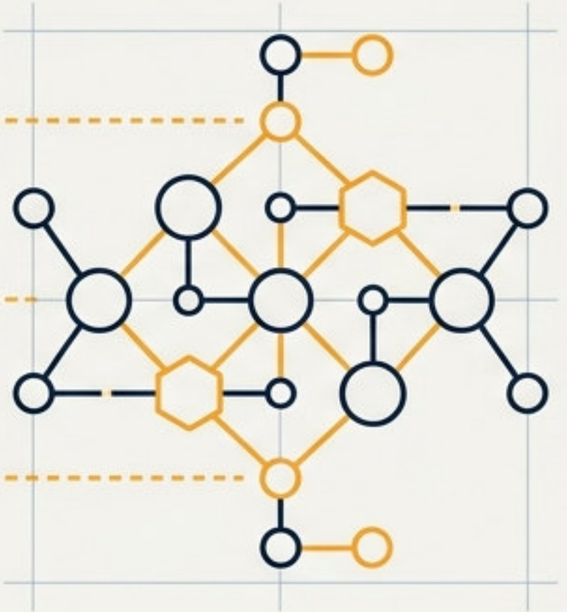
## The Stressor



Microgravity unloading drives osteoclast-biased bone remodelling.

- Result: ~1–2% areal bone mineral density (BMD) loss per month in astronauts.

## The Overlap with Ageing



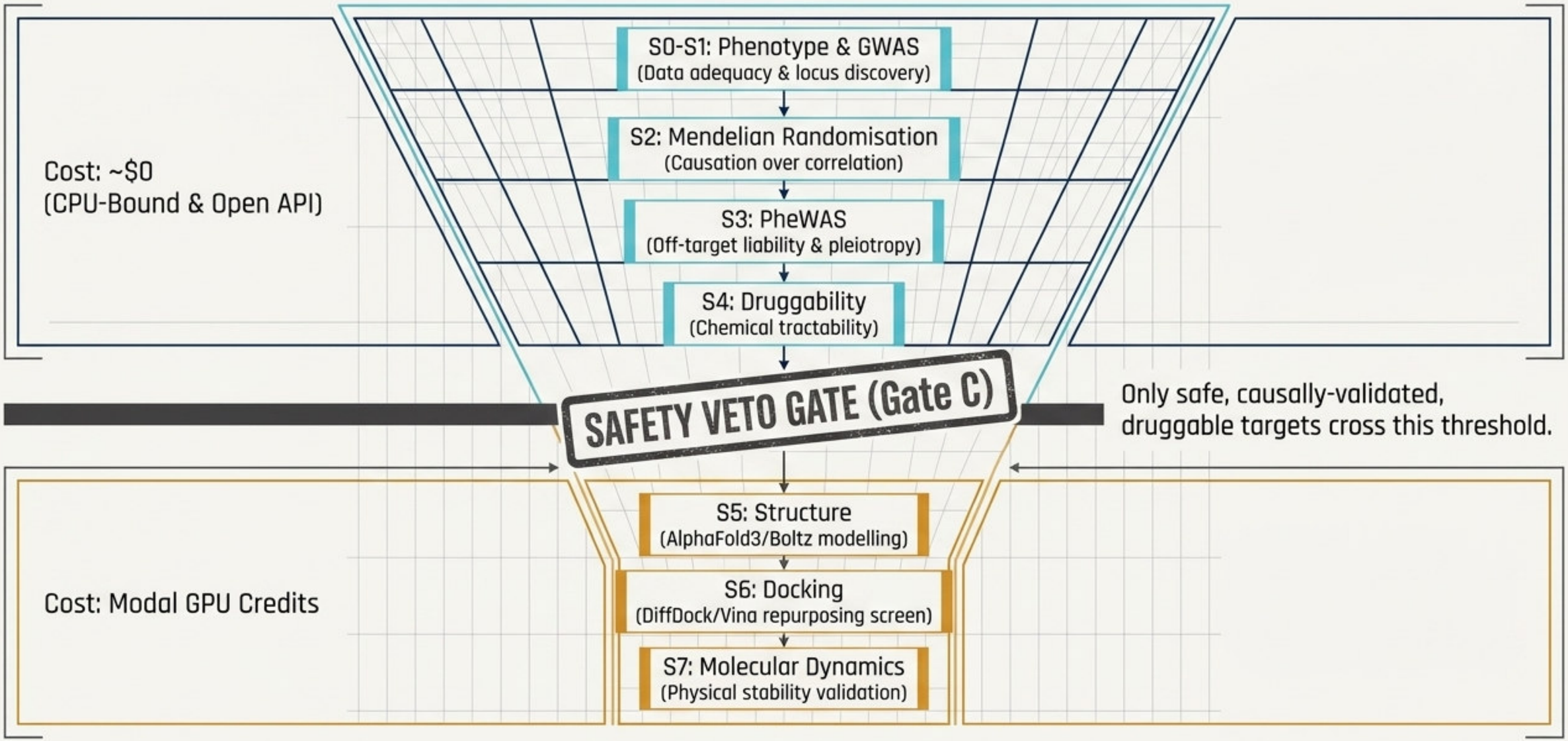
This aerospace stressor maps directly to terrestrial ageing hallmarks:

- Stem-cell exhaustion (MSC to osteoblast)
- Cellular senescence (osteocyte SASP)
- Altered intercellular communication (RANKL/WNT)

**The Mission Objective:** Mine human genetic variation to discover protective alleles, causally validate their effects, and repurpose small molecules to safely mimic their protection.

NotebookLM

# The Pipeline Architecture: A Cheap-to-Expensive Filtration Funnel




NotebookLM

# Strict Governance and Provenance by Default

## Agentic Harness


### Pre-registration

All statistical thresholds and multiplicity denominators (e.g., N_targets = 33) are frozen in a version-controlled params.yaml prior to execution. No relaxing rules for borderline hits.

### Contract-First Handoffs

Agents communicate exclusively through schema-validated JSON envelopes. No free-text guessing. Every stage is swappable and re-runnable.

### Immutable Provenance

Every artifact generates a run-manifest tracking database versions, API endpoints, random seeds, and SHA-256 hashes.

NotebookLM

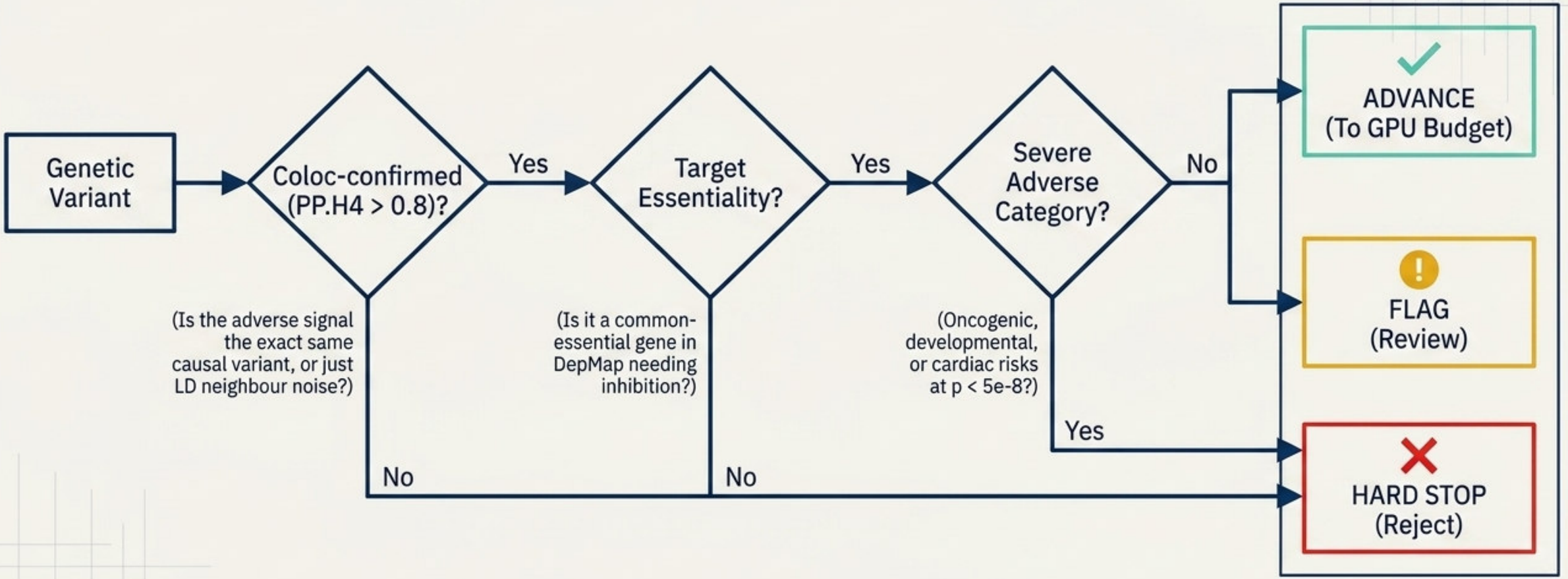
# Phase 1: The Genetics Engine & Target Discovery

Using heel estimated BMD (N=426,824) as the primary exposure to find independent instruments.

| Classic Mendelian Randomisation  | Drug-Target (cis-MR)  |
|---|---|
| The Core Question | The Core Question |
| Does a modifiable exposure (trait) cause the outcome? | Does pharmacologically modulating a specific target change the outcome? |
| Biological Precision | Biological Precision |
| **Instruments:** Genome-wide SNPs.<br>**Biology:** Distal, prone to pleiotropy.<br>**Use:** Triangulating biological pathways. | **Instruments:** cis-QTL variants only.<br>**Biology:** Proximal (variant → protein → phenotype).<br>**Use:** The primary discovery engine for druggability. |

NotebookLM

# The Crucial Safety Gate: Phenome-Wide Liability Screening




Only targets clearing these phenome-wide hurdles unlock the structural GPU budget.

NotebookLM

# Pilot Execution: The Genetics Verdict

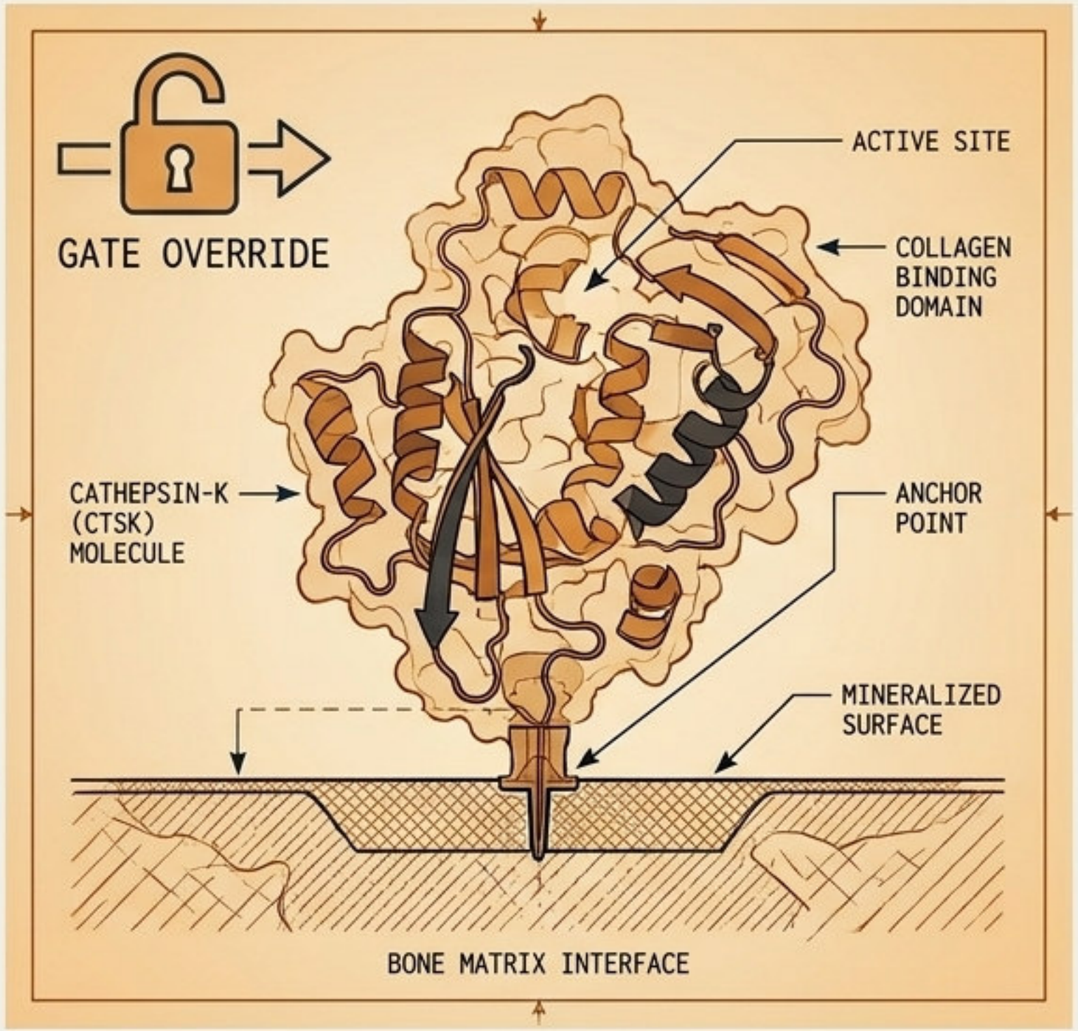
```
INPUT: 33 FROZEN DRUGGABLE TARGETS
STATUS: 26 lacked reachable large-N cis-eQTLs.
STATUS: 7 testable targets processed.
EVALUATING CAUSALITY GATE...
[PALE RED HIGHLIGHT RESULT: All 7 failed strict causality (Bonferroni α = 1.52e-03; coloc.abf PP.H4 < 0.8).
```


0 / 33 Targets Advanced

- **The Validation:** The eBMD→fracture positive control passed strongly (p = 1.85e-44), proving the MR machinery works perfectly.
- **The Conclusion:** A triumphant instrument-limited negative. The funnel successfully rejected weak signals rather than relaxing pre-registered thresholds to invent a false positive.

NotebookLM

# The CTSK Anchor: Advancing the Structural Engine




## The Strategy

To validate Stages 5-7 (Structure, Docking, MD) independently of the genetics negative, the pipeline advanced Cathepsin-K (CTSK) as a pre-specified small-molecule anchor.

## Biological Grounding

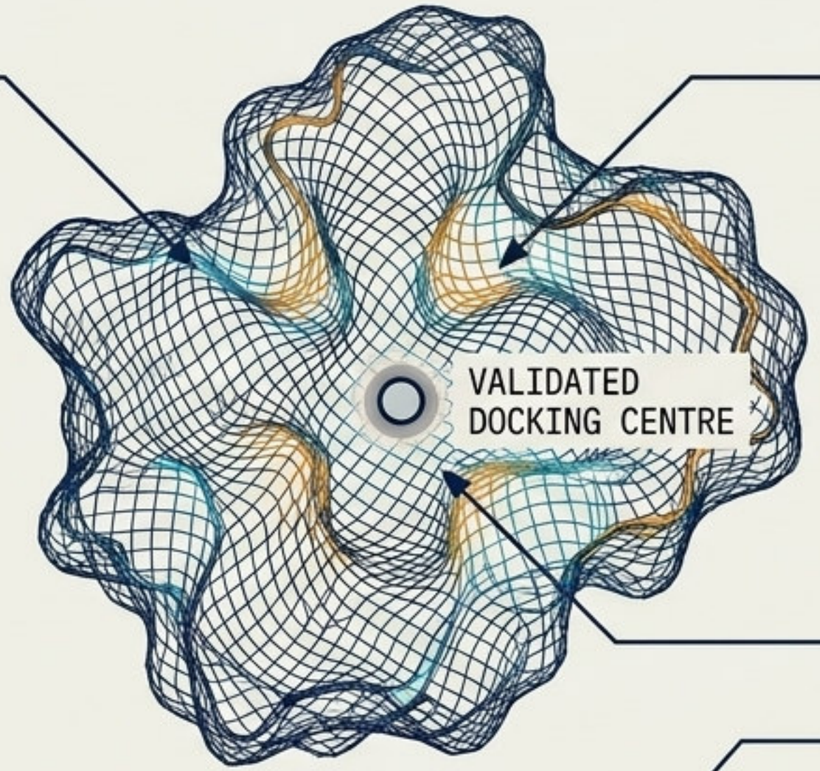
CTSK is the dominant osteoclast collagenase. Human loss-of-function causes **pycnodysostosis** (high bone density), explicitly validating inhibition as the required therapeutic direction to counteract spaceflight bone loss.

NotebookLM

# Structural Preparation & Docking Consensus

## Structural Security

Target structure confirmed via 5TDI (X-ray 1.4 Å) and AF2 model. Pocket confidence is exceptionally high (pLDDT 97.56). Sub-Ångström concordance (0.27 Å RMSD) between experimental and predicted pockets.




## Docking Gate (Gate D)

Evaluated using a two-method consensus (Vina CPU + Boltz-2 GPU).

## Redock Control

Odanacatib control successfully redocked at 1.026 Å (well under the strict < 2.0 Å requirement).

~1,155 library compounds

24-compound shortlist

NotebookLM

# Molecular Dynamics Verdict: Designing Away from Liability

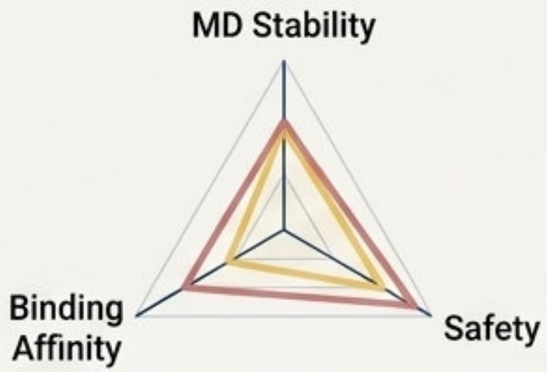
**The CV Caveat:** Odanacatib (control) was flagged at Gate C for realised **Phase-3 cardiovascular/stroke liability**. The **directive: find a structurally distinct alternative.**

Gate-C FLAG: CV Liability

## Odanacatib (The Control)

Confirms the MD protocol.

RMSD plateau 1.60 Å



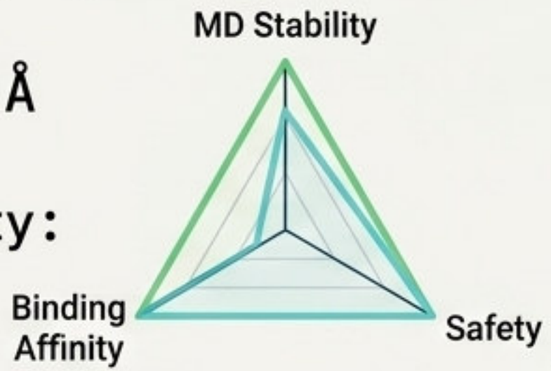

Chemotype-Distinct

## CHEMBL2164674 (The Lead)

Pipeline identified a physically stable candidate avoiding the control’s liability space.

RMSD plateau 0.99 Å

Tanimoto similarity: 0.165




NotebookLM

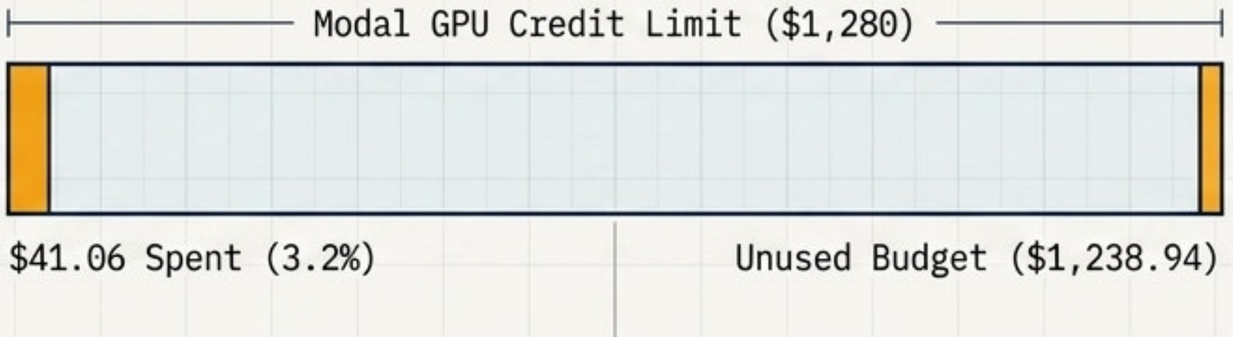
# High-Value Filtration: The $41.06 Blueprint

**$41.06** Total Compute Spend

## Compute Economy

Stages 0 through 4 (Genetics, Safety, Druggability) utilised CPU and open APIs, costing $0.00.

— CPU focus




## Strategic Spend

GPU compute was exclusively unleashed on pre-validated survivors. Stage 6 (Docking) utilised $2.02. Stage 7 (MD) utilised $39.04.

— GPU focus

## Total ROI

A causally-assessed genome, phenome-wide safety screen, and MD-validated compound shortlist executed for just 3.2% of the total available compute budget.

NotebookLM

## The Artifact

The Spacefarer Phenome provides a fully pre-registered, contract-driven, and safety-vetoed methodology for discovering physiological countermeasures.

Apache-2.0 (Code)

CC-BY-4.0 (Data/Report)

## Open Science

Fully auditable via immutable SHA-256 run-manifests, guarding against subjective post-hoc adjustments.

## The Future

A scalable, AI-orchestrated foundation ready to map the genomic frontier of human resilience in space.

NotebookLM