Deciphering Dopaminergic Heterogeneity: A GPU-Accelerated Evolutionary Framework Identifies Distinct Metabolic, Fibrotic, and Iron-Toxic States

Sam Neaves – UK Dementia Research Institute (UK DRI), Cardiff University

1. The "Mixed Bag" Challenge

Parkinson's Disease (PD) is asynchronous. At any moment, the substantia nigra contains a mixture of healthy, compensating, and dying cells. Standard clustering forces cells into mutually exclusive groups based on **Global Similarity**, obscuring transitional failure modes.

The Objective

To identify **robust, non-overlapping combinatorial rules** that define distinct cell states without prior biological assumptions.

2. The Solution: GPU-Accelerated DSSD

Method: Diverse Subgroup Set Discovery

We developed a custom Parallel Genetic Algorithm (GA) on **DAWN HPC**.

- **Glass Box AI:** Unlike Neural Networks, the GA produces logic rules.
- **Set Optimization:** Optimizes a **set** of rules simultaneously, penalizing overlap.

Rigorous Validation

- **Permutation Test:** 1,000 concurrent GPU searches ($p < 0.001$).
- **Shadow Run:** Blinded re-discovery to prove exhaustiveness.

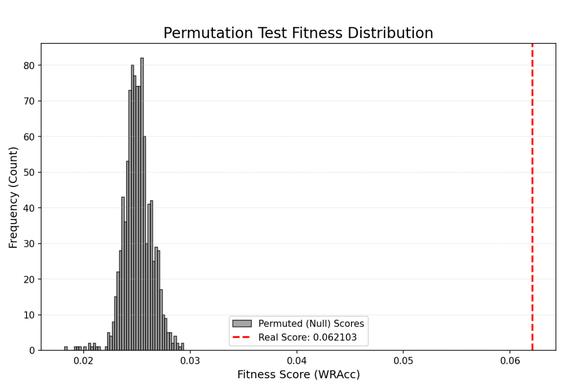

Figure 1: Discovery Signal ($p < 0.001$).

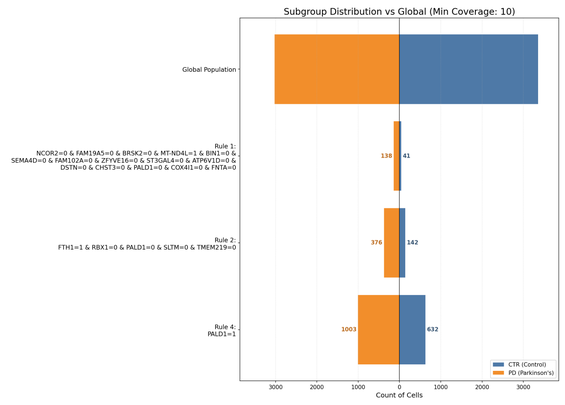

Figure 2: Subgroup Composition.

3. Discovery: A Triple Dissociation

The framework decomposed the PD population into three mutually exclusive "Attractor States."

Rule 1: Metabolic Compensation

Loss of nuclear repressor *NCOR2* triggers massive upregulation of Mitochondria. **Interpretation:** Cells fighting bioenergetic collapse.

Rule 2: Iron-Toxic Stress

Defined by Ferritin (*FTH1*) and Alpha-Crystallin B (*CRYAB*), a chaperone that prevents aggregation. **Interpretation:** Cells managing toxic iron accumulation.

Rule 4: Fibrotic Senescence

Enriched for *LINGO1* (blocks repair) and *COL27A1* (scarring). **Interpretation:** A terminal state of irreversible fibrosis.

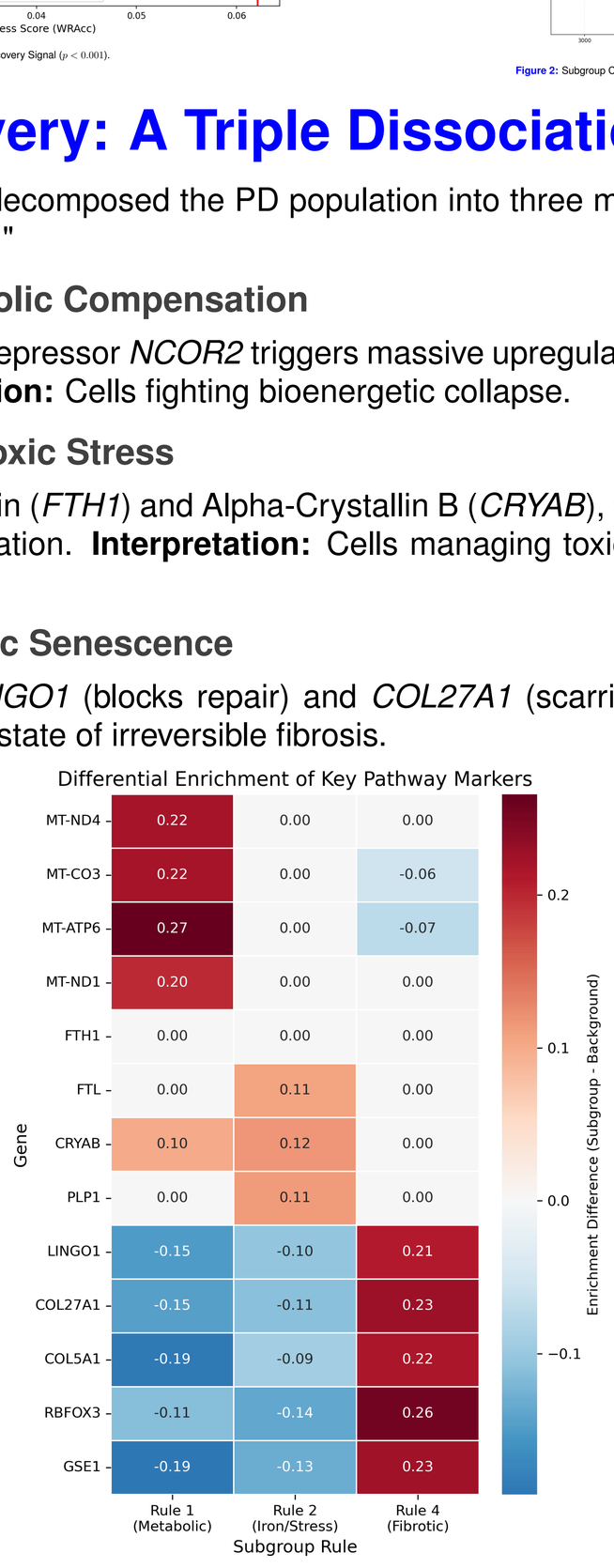

Figure 3: Differential Enrichment Heatmap. Note the "staircase" pattern: Mitochondrial genes (Top Left) vs Fibrotic genes (Bottom Right).

4. Validation: Biological Ground Truth

Validation against the Gene Ontology database confirmed the biological coherence of the discovered states.

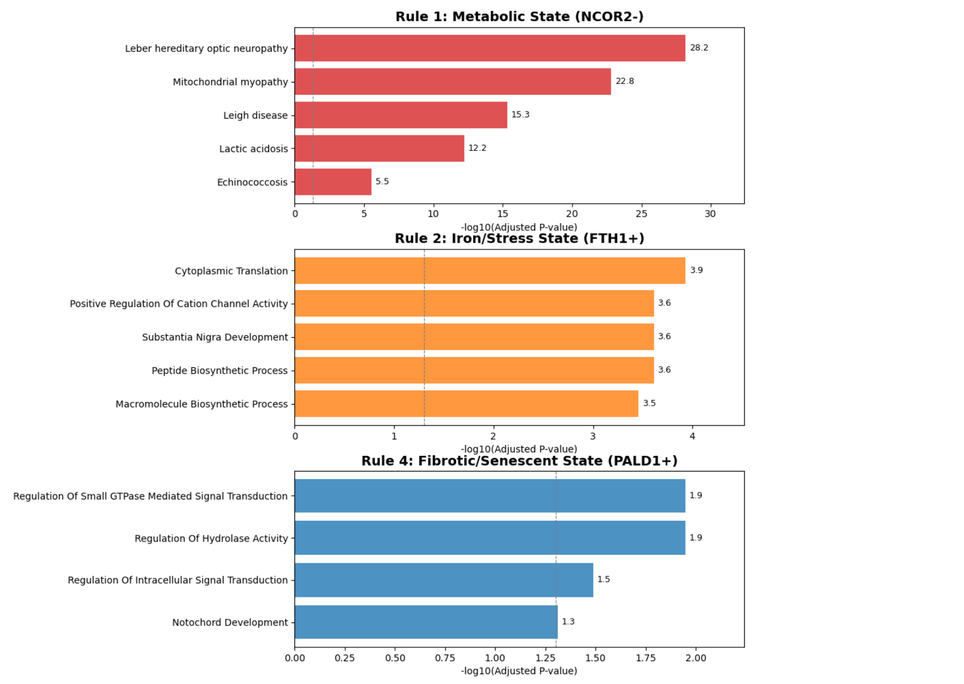

Figure 4: GO analysis confirms Rule 1 is Mitochondrial ($p < 10^{-20}$), Rule 2 is Iron/Translation, and Rule 4 is Fibrosis.

5. The "Trajectory of Failure"

We calculated the transcriptomic distance of each subgroup from a healthy control baseline to infer chronological progression.

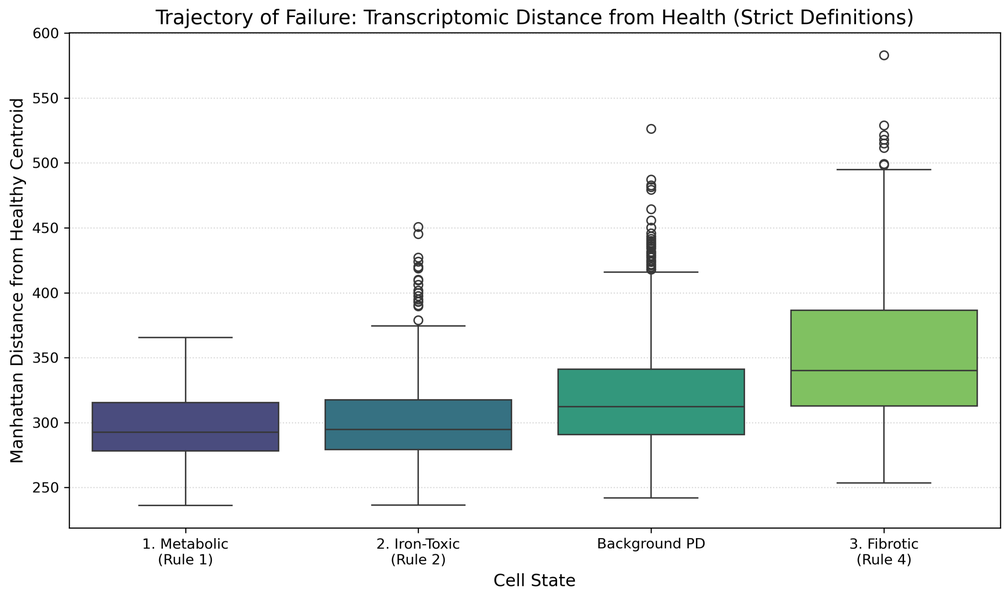

Figure 5: The Unified Model. Cells enter early stress states (Metabolic/Iron) before collapsing into the terminal Fibrotic state.

6. Impact: Functional States vs. Cell Identity

Standard clustering groups cells by global transcriptomic identity. Our framework reveals states that are **orthogonal** to this identity.

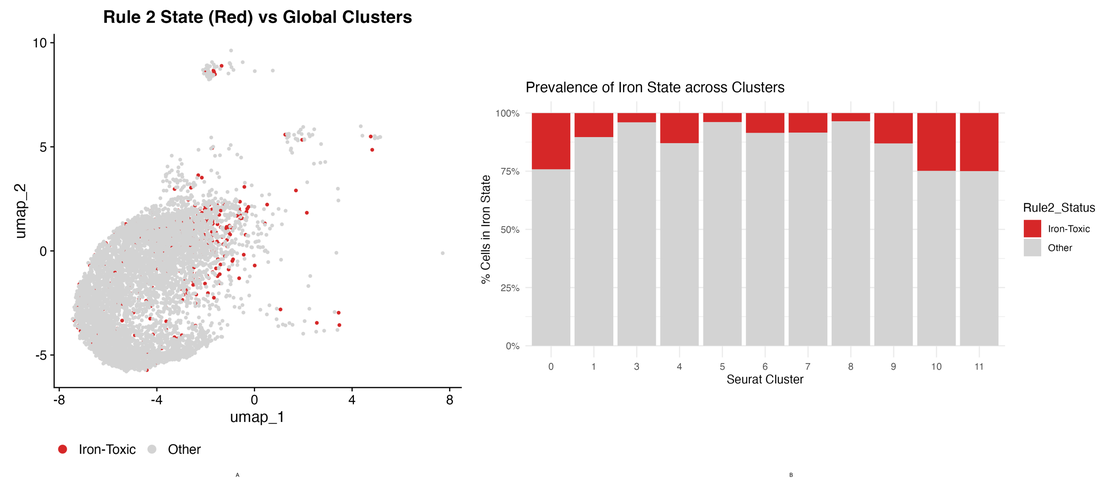

Figure 6: The "Iron-Toxic" state (Red) cuts across multiple standard clusters (Grey), proving it is a functional state acquired by various lineages, not a static cell type.

Conclusion & Precision Medicine

We have mapped the landscape of PD heterogeneity into three actionable states. This sheds light on why "one-size-fits-all" trials might ofail.

1. **Metabolic (Early):** Target for Mitochondrial Co-factors.
2. **Iron-Toxic (Early):** Target for Iron Chelation.
3. **Fibrotic (Terminal):** Target for Senolytics / Anti-LINGO1 therapies.