TOPOLOGICAL SIMPLIFICATION AND LOSS OF REGULATORY COMPLEXITY IN THE DNA REPAIR PATHWAY OF ACUTE LYMPHOBLASTIC LEUKEMIA

A. Mustafa[1]

1.MBBS, Faculty of Medicine, Al-Neelain University, Khartoum, Sudan.

INTRODUCTION

The DNA repair pathway is a highly coordinated, multilayered system of interacting genes, in which robust regulatory loops ensure genomic stability and precise response to DNA damage signals. Using Sheaf-theoretic spectral analysis and Hodge-Helmholtz decomposition, this study **aim** to investigate the disruption of DNA repair pathway complexity in acute lymphoblastic leukemia (ALL).

METHOD

Gene expression profiles for ALL were obtained from the GDC MP2PRT-ALL cohort, and control samples were sourced from the GTEx whole blood data. A curated gene set (>140 genes) defining the DNA repair pathway was retrieved from the Molecular Signatures Database (MSigDB). Functional coexpression networks were constructed, and a Hodge-Helmholtz decomposition was applied alongside Sheaf Laplacian spectral analysis to investigate the loss of complexity. Systemic network inconsistency was quantified via the H^1 harmonic residual energy, representing the residual harmonic 1-form component that cannot be reduced to a global gradient. The significance of the result was tested using a nonparametric permutation test.

RESULTS

The DNA repair pathway in ALL exhibits about a 52% reduction in H^1 obstruction energy relative to the normal state, indicating a loss of cyclic, nongradient correlation structure. Because this energy quantifies the residual harmonic 1-form component, its decline suggests the DNA repair pathway in ALL is less topologically frustrated and behaves more like a pure gradient flow. Biologically, this may indicate that the DNA repair pathway in ALL has lost complex, stabilizing regulatory loops, collapsing into a simpler hierarchical architecture. A permutation test shows that ALL exhibits a significantly lower H^1 energy than control (observed $\Delta=-0.0525$, empirical $P<0.0001$). Also, spectral analysis of the Sheaf Laplacian corroborates this finding. Here, the ALL network has a reduced trace (0.152 vs 0.331) and a smaller spectral gap, both hallmarks of a system with fewer independent constraints. Edge-wise harmonic residuals further pinpoint *RNMT*, *TP53*, *AAAS*, and *SUPT5H* as the top carriers of residual inconsistency in ALL.

CONCLUSIONS

The DNA repair pathway in ALL exhibits topological simplification marked by reduced H^1 harmonic energy and a loss of cyclic regulatory structures.

TABLES AND FIGURES

1. Data Preprocessing & Exploratory Analysis

i. Handle infinite values & remove 0-variance genes
ii. Generate baseline hierarchical clustering map
iii. Gene-wise Z-score normalization
iv. Hierarchical clustering heatmap
v. Global correlation thresholding ($|r| > 0.5$)
vi. Extract a shared biological network topology G(Vertices, Edges)

2A. Sheaf Laplacian Construction

i. Compute correlation matrices and apply Fisher transformation for variance stabilization
ii. Assign edge weights from stabilized correlations
iii. Compute corrected sheaf Laplacian ($L=d^{T}d$)
iv. Normalize by number of edges.
v. Eigenvalue decomposition, then compare Normal vs ALL spectra (trace, eigen ratio)
vi. Diffusion dynamics (heat kernel trace) analysis

2B. Hodge Residual Analysis

i. Build Coboundary Matrix B.
ii. Project correlations via Least Squares
iii. Compute H1 Obstruction energy from residuals
iv. Quantify global structural frustration ratio

3. Statistical Validation

i. Permutation testing (500 iterations)
ii. Null distribution
iii. Empirical P-value

Figure 1: Methodology Flow Diagram

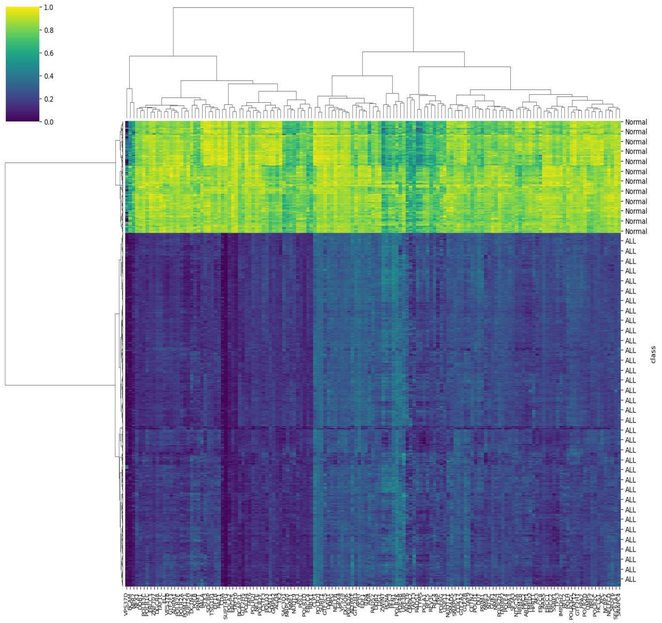

Figure 2: Heatmap of DNA Repair Pathway Genes in ALL vs Normal

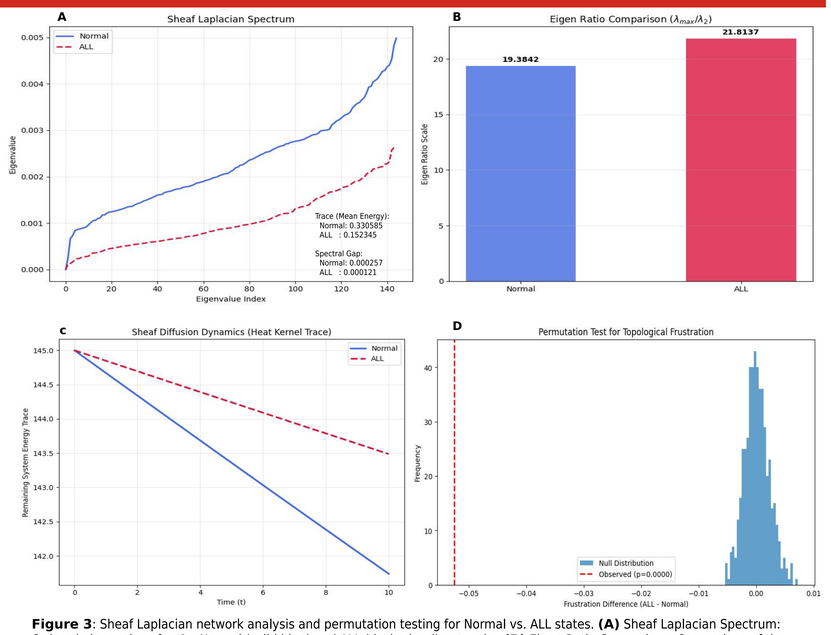

Figure 3: Sheaf Laplacian network analysis and permutation testing for Normal vs. ALL states. **(A)** Sheaf Laplacian Spectrum: Ordered eigenvalues for the Normal (solid blue) and ALL (dashed red) networks. **(B)** Eigen Ratio Comparison: Comparison of the maximum to second-smallest eigenvalue ratio, which serves as an indicator of synchronization capability. The ALL state exhibits a higher ratio than the Normal state, which reflects reduced capacity for global synchronization in ALL network compared to Normal. **(C)** Sheaf Diffusion Dynamics: Heat kernel trace showing the dissipation of system energy over a continuous time horizon (t). The faster decay in the Normal network indicates a faster diffusion rate relative to the ALL network. **(D)** Permutation Test for Topological Frustration: Empirical validation of the difference in Hodge residuals. The blue histogram displays the null distribution of frustration differences over 500 permutations, while the red dashed line marks the observed difference ($p = 0.0000$), indicating that the measured increase in topological frustration for the ALL network is statistically significant.

CONTACT INFORMATION

Alamin Mustafa, MBBS, Faculty of Medicine, Al-Neelain University, Khartoum, Sudan. Email: alamin900005@gmail.com

REFERENCES

Tahaei, M. S., Jalili, M., & Knyazeva, M. G. (2012). Synchronizability of EEG-based functional networks in early Alzheimer's disease. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 20(5), 636-641.

Su, Z., Tong, Y., & Wei, G. W. (2024). Hodge decomposition of single-cell RNA velocity. Journal of chemical information and modeling, 64(8), 3558-3568.

Bodnar, C., Di Giovanni, F., Chamberlain, B., Lio, P., & Bronstein, M. (2022). Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in gnns. Advances in Neural Information Processing Systems, 35, 18527-18541.

Merris, R. (1998). Laplacian graph eigenvectors. Linear algebra and its applications, 278(1-3), 221-236.

Xiao B, Hancock ER, Wilson RC. Graph characteristics from the heat kernel trace. Pattern Recognition. 2009 Nov 1;42(11):2589-606.