TOPOLOGICAL RECONSTRUCTION OF DNA REPAIR SIGNATURES AND SURVIVAL DIVERGENCE IN ACUTE MYELOID LEUKEMIA

A. Mustafa[1]

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

INTRODUCTION

Genomic instability is a hallmark of acute myeloid leukemia (AML), where abnormal DNA repair processes lead to chemoresistance and poor survival rates. The nonlinear, high-dimensional nature of DNA repair gene interactions underlying prognostic divergence are often missed by conventional survival models. This study **aims** to identify survival-divergent clusters linked to overall survival (OS), this work used a topological mapper to reconstruct the nonlinear geometric manifold of DNA repair dysregulation in AML.

METHOD

The Cancer Genome Atlas (TCGA) AML cohort with complete survival time and vital status data was analyzed (n=132). The Molecular Signature Database (MSigDB) provided a curated gene set (>140 genes) for the DNA repair pathway. Gene expression heat mapping was used for the initial visualization. Principal component analysis (PCA) reduced the dimensionality and produced a bivariate lens for the mapper algorithm. Density-based spatial clustering of applications with noise (DBSCAN) was used as the local clustering algorithm within a cubical cover (10 intervals, 30% overlap). Nodes were colored by survival time and vital status to visualize survival-associated topological stratification.

RESULTS

The Mapper graph demonstrated a densely connected blue core comprising predominantly alive patients with longer OS, representing a cohesive DNA repair expression state. In contrast, yellow nodes were sparsely distributed at the periphery, forming distinct branches enriched for deceased patients with shorter survival. Genes implicated in the DNA repair pathway, including cytidine deaminase (*CDA*) and phosphodiesterase 6G (*PDE6G*), demonstrated higher expression within poor-survival branches, indicating topological separation from the central survival-dominant manifold. Transitional nodes linking the core to peripheral branches suggested intermediate genomic states preceding poor outcomes.

CONCLUSIONS

This work bridges high-dimensional DNA repair gene expression with clinical survival divergence in AML. By identifying survival-segregated topological structures, this framework provides a visual and mechanistic basis for risk stratification and for optimizing DNA damage response-targeted therapies in a survival-informed manner.

TABLES AND FIGURES

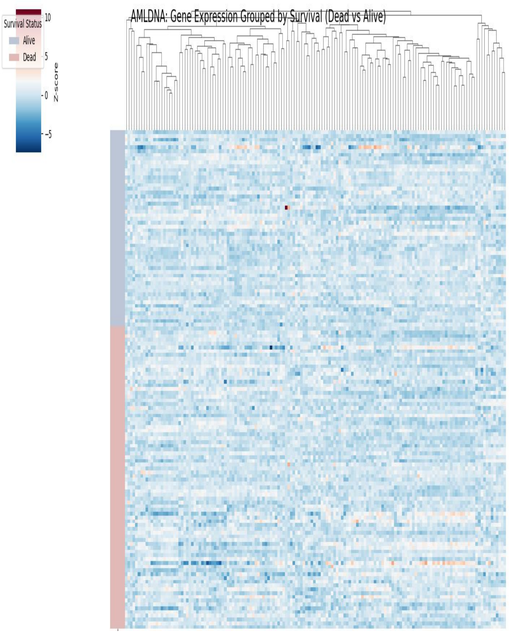

Figure 1: This heat-map shows the expression of DNA repair pathway genes in alive vs dead AML cohort.

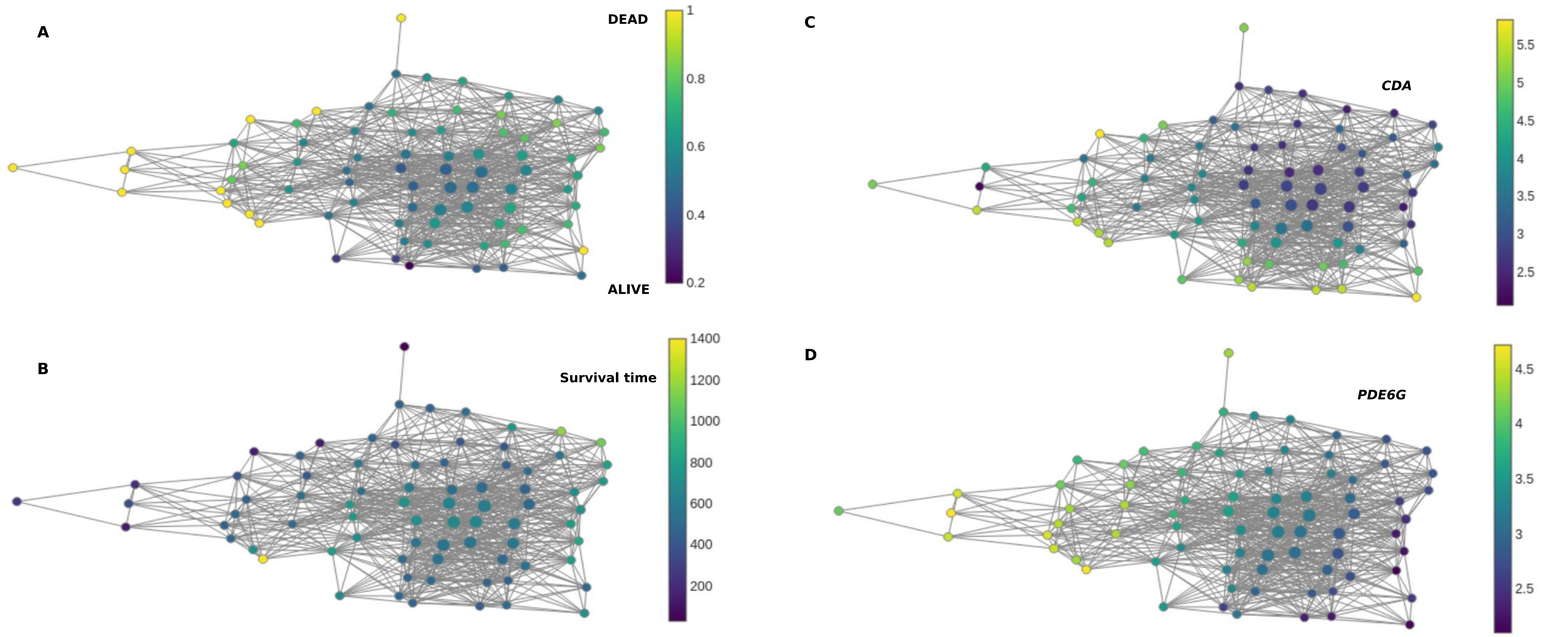

Figure 2: Topological Mapper graph analysis of patient survival and gene expression profiles. **(A)** Patient Survival Status: Distribution of survival status across the network. A dense central core consists predominantly of surviving patients (darker nodes), while sparse peripheral branches are heavily enriched with deceased patients (yellow nodes). **(B)** Overall Survival Time: Longitudinal mapping of survival duration (up to 1400 days). Longer overall survival times cluster within the central cohesive core, while shorter survival times align with the extended peripheral structures. **(C)** *CDA* Gene Expression: Elevated expression levels of *CDA* (yellow/green nodes) are localized within the peripheral, poor-survival branches, demonstrating topological separation from the central core. **(D)** *PDE6G* Gene Expression: Higher expression levels of *PDE6G* (yellow/green nodes) mirror the distribution of *CDA*, selectively localizing to the poor-outcome peripheral branches and indicating an intermediate-to-poor genomic transition state.

CONTACT INFORMATION

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

REFERENCES

Madukpe VN, Ugoala BC, Zulkepli NF. A comprehensive review of the mapper algorithm, a topological data analysis technique, and its applications across various fields (2007–2025). International Journal of Data Science and Analytics. 2026 Jun;21(1):56.

Haşegan D, Geniesse C, Chowdhury S, Saggar M. Deconstructing the Mapper algorithm to extract richer topological and temporal features from functional neuroimaging data. Network Neuroscience. 2024 Dec 10;8(4):1355-82.

Carlsson G. Topological pattern recognition for point cloud data. Acta Numerica. 2014 May;23:289-368.