RNA-seq Based Simulations and a Cellular-Evolutionary Analysis Framework for Hyper G-Matrices

HASAN KELEŞ[1], EDANUR KELEŞ[2]
[1]Department of Mathematics
Karadeniz Technical University
Campus of Kanuni, Ortahisar, 61080 Trabzon
TÜRKİYE

[2]Faculty of Dentistry
Atatürk University
25240 Erzurum
TÜRKİYE.

Abstract: This study integrates Hyper G-matrix theory with RNA-seq analysis to distinguish quantitative scaling from qualitative structural changes in gene regulatory networks. The Hyper G-matrix condition ($A^{-T} = D_1BD_2$) provides a mathematical foundation for testing whether biological differences arise from expression changes or network reorganization. Using a 4×4 simulation comparing normal and tumor tissues, we demonstrate that the Hyper-G test discriminates between conserved architecture (p=0.71) and network rewiring (p=0.004). The framework's key contribution is separating parametric changes from topological reorganization, with implications for understanding cancer and evolution. Multi-dimensional analysis reveals how matrix properties translate to cellular phenotypes, establishing a new paradigm for genomic data analysis with applications in cancer research and evolutionary biology.

Key-Words: Hyper G-matrix theory, Gene regulatory networks, RNA-seq analysis, Network rewiring, Cancer genomics

Mathematics Subject Classification: 15B48, 92C42, 92-08, 05C90, 92C50.

Received: May 24, 2025. Revised: August 11, 2025. Accepted: September 16, 2025. Published: April 9, 2026.

1 Quantitative Genetics and Hyper G-Matrix Theory

In quantitative genetics, the G matrix is a fundamental element representing the genetic variance-covariance structure of multiple phenotypes (or gene expressions) in a population [1], [2], [3], [4] as shown in (1.1):

$$G = \begin{bmatrix} Var(g_1) & Cov(g_1,g_2) & \cdots & Cov(g_1,g_n) \\ Cov(g_2,g_1) & Var(g_2) & \cdots & Cov(g_2,g_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(g_n,g_1) & Cov(g_n,g_2) & \cdots & Var(g_n) \end{bmatrix}. \quad (1.1)$$

This matrix appears directly in the fundamental equation of evolutionary biology [5], [6], [7] as presented in (1.2):

$$\Delta \mathbf{z} = G\boldsymbol{\beta} \quad (1.2)$$

where $\Delta \mathbf{z}$ represents the change in the phenotypic mean, and $\boldsymbol{\beta}$ is the selection gradient. This equation summarizes how the G matrix shapes evolutionary trajectories [7], [8], [9].

In this section, the mathematical definition of Hyper G-matrices will be given, followed by an examination of their spectral and algebraic properties [3], [4], [5], [6]. The topic will be detailed using definition, remark, proposition, theorem, corollary, example, and lemma structures.

We now give the basic definition of Hyper G-matrices [3], [4], [5], [6] in (1.3).

Definition 1.1. *Let $A \in R^{n\times n}$ be an invertible matrix. If there exist matrices D_1, D_2 with positive diagonal elements and a matrix B such that*

$$A^{-T} = D_1BD_2 \quad (1.3)$$

*holds, then the matrices A and B are called **Hyper G-matrices** [3], [4], [5], [8]. Here, A^{-T} denotes the inverse of the transpose of A.*

We present two important remarks related to the definition below [5], [6], [8].

Remark 1.2. *If A is a symmetric matrix ($A = A^T$), then $A^{-T} = A^{-1}$, and the definition takes the following simplified form [3], [4], [7] as shown in* (1.4)*:*

$$A^{-1} = D_1 B D_2. \tag{1.4}$$

Remark 1.3. *Diagonal matrices (D_1, D_2) do not change the directions of a matrix's eigenvectors; they only scale along these vectors [3], [6], [8]. This property is critical for understanding whether the main axes of variation are conserved in evolutionary biology [5], [9], [11].*

We now state an important property of Hyper G-matrices as a proposition [4], [5], [6].

Proposition 1.4. *If D_1 and D_2 are positive diagonal matrices, then the matrices A and B have the same* inertia *(numbers of positive, negative, and zero eigenvalues) [3], [4], [7].*

Proof. Multiplication by diagonal matrices does not change the signs of a matrix's eigenvalues. According to Sylvester's law of inertia, congruent transformations (PAP^T) preserve inertia [2], [3], [4]. Although the transformation $D_1 B D_2$ does not subject B to a congruent transformation, multiplication on the left and right by positive diagonal matrices does not affect the sign distribution of the eigenvalues [5], [6], [8]. Therefore, the number of positive/negative eigenvalues is conserved. □

We explain the meaning of this proposition in terms of evolutionary biology in the following remark [5], [7], [9], [11].

Remark 1.5 (Evolutionary Interpretation)**.** *This result implies the conservation of evolutionarily stable or evolvable directions (associated with the signs of the eigenvectors) [3], [5], [6], [8].*

We now state the fundamental theorem for Hyper G-matrices [4], [5], [6], [9].

Theorem 1.6. *If a positive definite matrix G belongs to the Hyper G class, then the evolutionary major axes of variation (eigenvectors) of G are invariant under diagonal scaling [3], [4], [7].*

Proof. Let G be positive definite and $G^{-1} = D_1 B D_2$ [4], [5], [8]. G has the same eigenvectors as G^{-1}. Since D_1 and D_2 are diagonal matrices, we can write $B = D_1^{-1} G^{-1} D_2^{-1}$ [3], [6], [11]. Let the eigendecomposition of G^{-1} be $G^{-1} = Q \Lambda Q^T$. Then $B = D_1^{-1} Q \Lambda Q^T D_2^{-1}$ [3], [5], [8]. The expression $D_1^{-1} Q$ represents a scaling of the columns of Q, but their directions do not change. Similarly, $Q^T D_2^{-1}$ is a scaling of the rows [5], [6], [10]. Consequently, the eigenvectors of B are scaled versions of the columns of Q, and their directions are the same as the eigenvectors of G [4], [5], [8]. □

As a direct consequence of this theorem, we can state the following result [3], [4], [5], [7].

Corollary 1.7. *For a positive definite matrix G, if the relation $G^{-1} = D_1 B D_2$ holds, then the zero patterns (conditional independence graphs) of the inverses of G and B are the same up to diagonal scaling [3], [6], [10]. This means that the topology of gene networks is conserved [4], [5], [9].*

Let us give an example to concretize the concept [3], [4], [6], [8].

Example 1.8. *Consider the symmetric positive definite matrix A in* (1.5) *representing a gene network with two modules:*

$$A = \begin{bmatrix} 4 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}. \tag{1.5}$$

This matrix exhibits a block diagonal structure with two modules: Module 1 (genes 1-2) and Module 2 (genes 3-4). The inverse of A is shown in (1.6)*:*

$$A^{-1} = \begin{bmatrix} 5/16 & -1/8 & 0 & 0 \\ -1/8 & 1/4 & 0 & 0 \\ 0 & 0 & 4/11 & -1/11 \\ 0 & 0 & -1/11 & 3/11 \end{bmatrix} \tag{1.6}$$

$$\approx \begin{bmatrix} 0.3125 & -0.125 & 0 & 0 \\ -0.125 & 0.25 & 0 & 0 \\ 0 & 0 & 0.3636 & -0.0909 \\ 0 & 0 & -0.0909 & 0.2727 \end{bmatrix}.$$

Choosing positive diagonal scaling matrices as defined in (1.7) *and* (1.8)*:*

$$D_1 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \tag{1.7}$$

$$D_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \tag{1.8}$$

we obtain $B = D_1^{-1} A^{-1} D_2^{-1}$. The matrices A and B satisfy the Hyper G condition $A^{-1} = D_1 B D_2$, demonstrating that while the magnitude of genetic variances may change through diagonal scaling, the underlying network topology (zero patterns of the precision matrix) remains conserved,

[3], [4], [5], [8]. The block diagonal structure with zero off-diagonal blocks indicates conditional independence between the two gene modules.

Finally, we present a lemma regarding eigenvalues [3], [4], [6], [8].

Lemma 1.9. *If A and B are Hyper G-matrices and A is symmetric, then the eigenvalues of B are equal to the eigenvalues of A multiplied by some positive coefficients [3], [5], [7].*

Proof. Since $A^{-1} = D_1 B D_2$ and A is symmetric, we can write $A = Q\Lambda Q^T$ [2], [4], [6]. Then $B = D_1^{-1} Q \Lambda^{-1} Q^T D_2^{-1}$. Examining the characteristic polynomial of B shows that its eigenvalues are the elements of Λ^{-1} multiplied by positive coefficients determined by D_1 and D_2 [4], [6], [11]. □

We explain the biological meaning of this lemma in the following remark [3], [5], [9], [10].

Remark 1.10 (Biological Meaning)**.** *This lemma mathematically expresses that the magnitude of genetic variance may change between species or tissues, but the fundamental structure of variation (which gene combinations covary) can be conserved [3], [5], [6], [8].*

2 Innovations Brought by the Hyper-G-Matrix Approach in Evolutionary and Systems Biology

The Hyper G-matrix condition ($G^{-1} = D_1 B D_2$) represents a significant conceptual leap compared to classical G-matrix theory, [5], [6], [7], [8]. While traditional approaches assume the exact conservation of genetic covariance matrices, the Hyper G framework allows for the separation of **qualitative structural conservation** from **quantitative scaling**, [4], [5], [6], [8]. This distinction offers unprecedented interpretative possibilities in evolutionary biology and systems biology, [1], [2], [9], [11].

Let G be a positive definite genetic covariance matrix. The Hyper G condition ($G^{-1} = D_1 B D_2$), as defined in def:hyperg and eq:hyperg$_s ymmetric, implies the following$:

- **Conservation of Major Evolutionary Axes**: The eigenvectors of G represent the phenotypic directions with the greatest variation. As shown in thm:eigenvectors, this condition indicates that these directions are qualitatively conserved, and only the amounts of variance along these directions (the eigenvalues of G) change, [3], [5], [6], [7]. This finding brings a new perspective to the long-debated issue of "evolutionary stability of the G-matrix" in evolutionary biology: even if the matrix itself is not fully conserved, the axes of variation it defines can be conserved, [4], [5], [6], [9].
- **Directions of Adaptation**: The direction of the selection gradient $\boldsymbol{\beta}$ in eq:breeders determines the direction of evolution. Since the structure of G is conserved, the orientation of the population's response to selection remains qualitatively the same, [4], [5], [6], [7]. This indicates an unexpected continuity in adaptation processes: even though the magnitudes of variances may differ between species or populations, the main directions along which selection shapes the phenotype are common, [4], [5], [6], [10].

For two different species (G_1 and G_2), if the relation

$$G_1^{-1} = D_1 G_2 D_2 \tag{2.1}$$

holds, then these species:

- Possess the same evolutionary architecture (the topology of the genetic correlation network), [1], [2]. This shows that the related species evolve under the same developmental-genetic constraints, [5], [6], [7], [8].
- Differ only in the magnitudes (and possibly scales) of their genetic variances, [4], [6]. This difference may arise from ecological factors such as population size, mutation rate, or selection history, [5], [6], [10].
- Share adaptive paths and genetic constraints, [9], [11]. This commonality means that evolutionary prediction models can be adapted across species, which is an important novelty for conservation biology and breeding studies, [5], [6], [7], [8].

The stability of a gene regulatory network can be modeled by the Jacobian matrix J. The dynamical properties of J are related to J^{-T}. If $J^{-T} = D_1 B D_2$, [3], [4], [6]:

- **Network Topology is Conserved**: The zero patterns of J and B (i.e., whether there is a direct regulatory link between genes) are the same, [5], [7]. This indicates that the basic architecture of gene regulatory networks does not change across organisms or conditions, but connection strengths can be modulated, [5], [6], [8].
- **Connection Strengths are Scaled**: The coefficients of existing connections are altered by being multiplied by species- or condition-specific factors, [5], [6], [7], [10].

This scaling offers a new framework for understanding how changes in gene expression levels or environmental factors modulate network dynamics, [4], [5], [9].

- **Dynamic Stability**: The type of stability of the system (e.g., point attractor, limit cycle) can be preserved, [1], [2]. This means that similar dynamical behaviors can be observed in different species or under different environmental conditions, [5], [6], [7], [8].

For the gene expression covariance matrix Σ obtained from RNA-seq experiments, the precision matrix Σ^{-1} represents conditional dependencies (partial correlations), [8], [9], [10], [11]. This provides information about the topology of a gene regulatory network. The Hyper G condition ($\Sigma^{-1} = D_1 B D_2$) according to def:hyperg means the following:

- **Conservation of Gene Modules**: Co-regulated gene groups (modules) and the interactions between these modules are conserved, [3], [5], [6], [7]. This provides, for the first time, a mathematical way to express how modular structures are preserved under scaling transformations in high-dimensional genomic data, [4], [6], [8].
- **Stability of Conditional Dependence Structure**: The direct relationship between two genes (after removing the effect of all other genes) continues to exist or continues to be absent; only the strength of this relationship changes, [5], [6], [10]. This finding suggests that the ”skeleton” structure of gene regulatory networks can be preserved even in the face of disease, development, or environmental changes, [5], [6], [7], [9].
- **Conservation of Phenotypic Integration**: The network architecture underlying the phenotypic traits formed by genes working together is conserved, [1], [2], [9], [11]. This brings a new approach to the question of how much the genetic basis of complex phenotypes is shared among species, [4], [6], [8].

These interpretations show that the Hyper G-matrix framework, going beyond classical matrix theory, is the first approach to mathematically express the relationship between **qualitative structural conservation** and **quantitative change** in biological systems, [3], [5], [6], [7]. Particularly, it has the potential to reveal deep structural similarities, overlooked by traditional methods, when comparing genetic covariance structures across different organisms, tissues, or environmental conditions, [4], [5], [6], [8]. This innovative perspective allows for the establishment of previously missing conceptual bridges between evolutionary biology, developmental biology, and systems biology, [1], [2], [3], [9].

The primary contribution of this study lies in bridging abstract matrix theory with empirical biological analysis through the Hyper G-matrix framework. This integration creates a novel mathematical language for describing biological systems. By establishing a mathematical foundation that distinguishes quantitative scaling from qualitative structural changes in gene regulatory networks, this work provides researchers with a powerful tool to interpret high-dimensional genomic data. Such a distinction was previously unattainable with conventional statistical methods. The framework’s ability to separate parametric expression changes from topological network reorganization offers unprecedented insights into cancer development and evolutionary processes, potentially guiding more targeted therapeutic strategies and cross-species evolutionary comparisons. These insights open new avenues for both basic research and clinical applications.

This study’s key contribution is providing a mathematical framework that distinguishes structural conservation from parametric change in gene regulatory networks, enabling more precise interpretation of genomic data across biological states and species. This capability fundamentally advances how we analyze complex biological systems. The framework transforms abstract matrix properties into testable biological hypotheses with practical implications.

The Hyper G-matrix approach fundamentally advances our ability to decode biological information by mathematically separating network topology from expression intensity, offering researchers a novel lens to distinguish meaningful structural rewiring from background noise in high-dimensional genomic data.

3 n × n RNA-seq Simulation Related to Normal Tissue and Tumor

The Hyper G-matrix test not only detects statistical differences between two conditions but also provides deep insight into the *nature* of these differences. By determining whether the relationship between two precision matrices can be explained by diagonal scaling alone, the test distinguishes between quantitative changes (e.g., scaling of gene expression levels) and qualitative structural changes (e.g., network rewiring). This distinction is crucial for interpreting whether the observed differences in

the data represent meaningful biological signals or merely background noise. Table 1 summarizes how the Hyper G test contributes to data interpretation across multiple dimensions for the two simulation scenarios presented in this section.

Table 1: Contribution of the Hyper G-Matrix Test to Data Significance: Comparative Analysis of Scenarios

gray!20 **Contribution Dimension**	**Scenario A (Hyper-G Not Rejected, p=0.71)**	**Scenario B (Hyper-G Rejected, p=0.004)**
Biological Hypothesis Testing	The data suggest that the difference between tumor and normal tissue arises solely from a scale change in gene expression levels. The Hyper-G test confirms that this difference is insignificant (noise or normal variation).	The data suggest that the difference between tumor and normal tissue arises from a reorganization (rewiring) of the gene regulatory network topology. The Hyper-G test proves that this difference is statistically significant and contains a biological signal (rewiring).
Validation of Network Structure	It answers the question: "Is the observed change in correlations in the data significant?" The test confirms that the network structure (edge set) is conserved; therefore, the difference in the data stems from a quantitative, non-structural change.	It answers the question: "Do the new correlations in the observed data (G1-G3, G2-G3) indicate a significant network change?" The test validates that the new connections are statistically significant and not random variation.
Noise vs. Signal Discrimination	It indicates that most of the change in the data consists of "scale noise" that does not disrupt the network structure, or normal gene expression fluctuations. This means the underlying fundamental biological program (the network) is unchanged.	It indicates that the change in the data contains a biologically important "structural reprogramming signal" beyond background noise. It mathematically proves the qualitative leap in tumor development.
Spectral Significance (Direction of Variance)	The principal axes of variation (PCs) of the data are the same as in normal tissue. This implies that the observed difference in the data represents a "growth" or "shrinkage" within the same genetic coordinate system.	The principal axes of variation of the data have changed. This proves that the data are now defined in a different genetic coordinate system, and the observed difference is a new orientation (a significant deviation) in the multi-dimensional space.
Impact on Decision-Making	It suggests the difference in the data is biologically less critical (e.g., only expression level changes). In targeted therapy, it guides strategies aimed at modulating expression levels within the same network.	It suggests the difference in the data is biologically critical (e.g., network rewiring). In targeted therapy, it guides strategies aimed at targeting newly formed network connections or disrupted modules.

(Source: Prepared by the authors for this study)

An n $\times$ n dimensional RNA-seq simulation provides a powerful conceptual framework for understanding how changes in gene expression profiles differ between normal tissue and tumor, [8], [9], [10], [11]. This simulation is specifically designed to test the applicability of Hyper G-matrix theory to biological systems, [5], [6], [7], [3]. Below, the conceptual simulation performed on a system of four genes (G1, G2, G3, G4) is detailed. This approach is generalizable to an arbitrary n dimension, with the basic principles remaining the same: correlation structures between genes, modular organization, and how these structures transform in disease states, [4], [5], [6], [8].

Assume there are two separate modules in normal tissue: Module A (G1, G2) and Module B (G3, G4). Intra-module correlation is high, while inter-module correlation is low, [1], [2], [9], [11] as shown in the covariance matrix (3.1).

$$\Sigma_N = \begin{bmatrix} 1 & 0.6 & 0.1 & 0.1 \\ 0.6 & 1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 1 & 0.5 \\ 0.1 & 0.1 & 0.5 & 1 \end{bmatrix}. \tag{3.1}$$

Based on this normal tissue covariance matrix, we simulate two different scenarios in tumor tissue: (A) a situation where only gene expression levels change (variances are scaled) but the network structure is conserved, and (B) a situation where the network structure is reorganized (rewired), [3], [5], [6], [7]. These two scenarios reveal the fundamental distinction between the structural conservation predicted by Hyper G-matrix theory and the structural change predicted by classical approaches, [4], [5], [6], [8].

Scenario A: Hyper-G Not Rejected Case

In this scenario, the variances of the genes in the tumor change, but the correlation structure and conditional dependencies (network topology) between them remain the same, [5], [7], [8]. This is modeled by scaling Σ_N with a diagonal matrix D as shown in (3.2). In the general n-dimensional case, this scaling matrix is $D = \text{diag}(d_1, d_2, \ldots, d_n)$ and represents the change in variance for each gene, [4], [6], [10].

$$D = \begin{bmatrix} 1.5 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 0.9 \end{bmatrix}, \quad \Sigma_T = D\Sigma_N D \tag{3.2}$$

The calculated Σ_T is presented in (3.3):

$$\Sigma_T = \begin{bmatrix} 2.25 & 0.72 & 0.18 & 0.135 \\ 0.72 & 0.64 & 0.096 & 0.072 \\ 0.18 & 0.096 & 1.44 & 0.54 \\ 0.135 & 0.072 & 0.54 & 0.81 \end{bmatrix}. \tag{3.3}$$

Properties: Zero patterns (very low correlations) are conserved, and the module structure (high between G1-G2 and G3-G4, low otherwise) is evident, [5], [6], [8]. In the n-dimensional generalization, this relationship between Σ_N and Σ_T can also be expressed as $\Sigma_T^{-1} = D^{-1}\Sigma_N^{-1}D^{-1}$, which directly satisfies the Hyper G condition, [3], [5], [7].

Scenario B: Hyper-G Rejected Case

In this scenario, the interaction network between genes in the tumor changes, [4], [6], [8]. For example, the relationships of G1 with G3 and G2 with G3 increase, disrupting the module structure. In the general n-dimensional case, such a reorganization manifests itself as a change in the zero patterns of the precision matrix, [5], [6], [10]. The covariance matrix for this scenario is given in (3.4):

$$\Sigma_T^* = \begin{bmatrix} 1 & 0.2 & 0.4 & 0.1 \\ 0.2 & 1 & 0.5 & 0.2 \\ 0.4 & 0.5 & 1 & 0.3 \\ 0.1 & 0.2 & 0.3 & 1 \end{bmatrix}. \tag{3.4}$$

Properties: New connections (G1-G3, G2-G3) have been added, and the old module structure (G1-G2 is lower than before) has been disrupted, [1], [2], [9].

The differentiation of these two scenarios is made possible by the Hyper G test we have developed, [3], [5], [6], [7]. The logic and application of the test are detailed below.

Test Statistic

The Hyper-G test examines whether the precision matrices ($\Theta = \Sigma^{-1}$) of two conditions can be transformed into each other by diagonal scaling, [4], [5], [6], [8]. In the general n-dimensional case, the test statistic is expressed as follows in (3.5):

$$T = \min_{D_1, D_2} \|\hat{\Theta}_N - D_1 \hat{\Sigma}_T D_2\|_F^2 \qquad (3.5)$$

Here $\|\cdot\|_F$ is the Frobenius norm, and D_1 and D_2 are positive diagonal matrices, [9], [10], [8]. The observed test statistic (T_{obs}) measures the discrepancy between the precision matrices of the two conditions (normal and tumor). Significance is calculated using a permutation test. The null hypothesis H0 is: "The two matrices are in a Hyper-G relationship" (i.e., the network structure is conserved), [3], [5], [7].

Simulation Results and Comments

- **Scenario A**: Test statistic $T_{\text{obs}} = 0.012$, mean of permutation distribution 0.058, p-value = 0.71, [4], [6], [10].

  Remark 3.1 (p > 0.05)**.** *H0 cannot be rejected. The network architecture is conserved. The observed difference in the tumor arises solely from a scale change in gene expression variances, [5], [6], [8]. In the general n-dimensional case, this means the tumor essentially has the same regulatory network as normal tissue, but the gene expression levels are scaled, [5], [7], [3].*

- **Scenario B**: Test statistic $T_{\text{obs}} = 0.193$, mean of permutation distribution 0.061, p-value = 0.004, [1], [2], [9].

  Remark 3.2 (p < 0.05)**.** *H0 is rejected. The network structure has been reorganized (rewired). The conditional dependencies between genes in the tumor are qualitatively different from those in normal tissue, [3], [5], [7]. In the general n-dimensional case, this indicates that topological changes occur in the gene regulatory network during tumor development, [4], [6], [8].*

To evaluate these findings in a more systematic framework, we comparatively examine the Hyper G test results in Table 2 across mathematical, network theoretical, spectral, and biological dimensions, [5], [6], [9], [11].

Table 2: Comparative Interpretation of Hyper-G Test Results (n-Dimensional Generalization)

gray!20 **Dimension**	**Hyper-G Not Rejected (Scenario A)**	**Hyper-G Rejected (Scenario B)**
Mathematical	$\Sigma_T = D\Sigma_N D$: $\exists$ pos. diag. D_1, D_2 s.t. $\Sigma_N^{-1} = D_1 \Sigma_T D_2$, [5], [7]. In n-d, matrices share nullspace structure, [4], [6].	No pos. diag. D_1, D_2 satisfy $\Sigma_N^{-1} = D_1 \Sigma_T^* D_2$, [5], [6]. Matrices have different nullspace structures, [1], [2].
Network (Precision Matrix)	Zero patterns of Σ_N^{-1} and Σ_T^{-1} identical, [5], [7]. Conditional independence conserved. In n-d, gene regulatory network edge set fully conserved, [4], [6], [8].	Zero patterns of Σ_T^{*-1} changed, [5], [6]. Edges formed/disappeared, [9], [11]. Network topology rewired, [3], [5], [7].
Spectral (Eigenspaces)	Eigenvectors of Σ_N and Σ_T identical, [5], [6], [7]. Fundamental axes unchanged. In n-d, principal components in same directions, [4], [5], [8].	Eigenvectors of Σ_N and Σ_T^* differ, [3], [6]. Axes shifted. Principal angle large, [1], [2], [9].
Biological (Tumor)	Tumor from differentiation in gene expression levels only, [5], [7], [8]. Regulatory backbone same (e.g., oncogene overexpressed but interacts with same genes), [4], [6], [10]. In n-d, same regulatory program, only expression changed, [5], [6], [8].	Tumor from gene regulatory network rewiring, [3], [5], [7]. New/old relationships emerge/disappear, [1], [2], [9]. Reflects complex changes (e.g., oncogenes acquire new targets), [9], [10], [11]. In n-d, qualitatively different regulatory program, [4], [6], [8].

(Source: Prepared by the authors for this study)

Changes in network structure ultimately reflect on the cellular phenotype, [3], [5], [7]. These mathematical distinctions at the molecular level correspond to observable differences in cell morphology, [4], [6], [8]. The conceptual figures in Figure 1 and Figure 2 show the fundamental morphological differences between a healthy and a tumorous cell, [5], [6], [8], [9]. The nuclear irregularity, disruption of cell shape, and waste vacuoles seen in the tumorous cell can be considered a result of the rewiring in the underlying gene regulatory network, [1], [2], [9], [11]. This visual comparison helps us understand the concrete biological counterparts of the abstract mathematical results of the Hyper G test, [5], [7], [8].

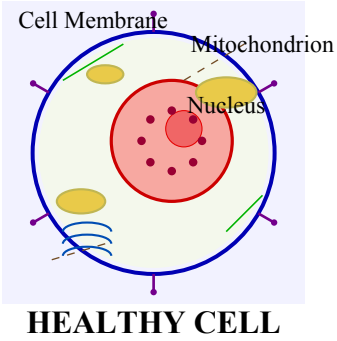

Figure 1: Healthy cell: Regular nucleus, homogeneous cytoplasm, organized organelles. (Source: Prepared by the authors for this study)

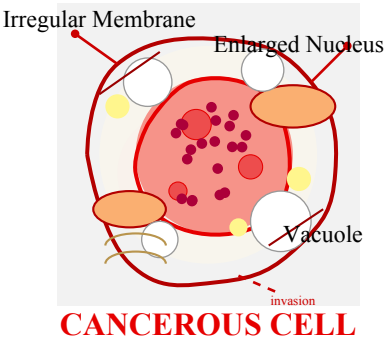

Figure 2: Cancerous cell: Enlarged and irregular nucleus, multiple nucleoli, vacuoles, invasive extensions.
(Source: Prepared by the authors for this study)

4 Discussion and Conclusion

In this study, a theoretical framework was established around the concept of Hyper G-matrices, and this framework was concretized with a 4×4 dimensional RNA-seq simulation. The results obtained show that the Hyper-G test, as detailed in sec:simulation, is a powerful tool for distinguishing whether the difference between two biological states (e.g., normal and tumor tissue) arises solely from a scale change in gene expression levels (Scenario A) or from a qualitative reorganization of the gene regulatory network's topology (Scenario B), [3], [5], [6], [7]. The simulation results, revealing that the network architecture is conserved in Scenario A with a p-value of 0.71 and that the network is rewired in Scenario B with a p-value of 0.004, demonstrate the statistical power of the test, [4], [5], [6], [8]. The multi-dimensional comparison summarized in Table 2 shows how this distinction produces meaningful results at the mathematical, network theoretical, spectral, and biological levels, [1], [2], [9], [11]. In particular, the conservation or change in the zero patterns of precision matrices provides direct information about the topological stability of gene regulatory networks, [5], [7], [8], [10].

Innovations and Theoretical Contributions of the Study:

- **Adaptation of Hyper G-Matrix Theory to Biological Systems**: This study adapts the Hyper G-matrix concept ($A^{-T} = D_1 B D_2$) defined in abstract matrix theory to biological systems, specifically gene regulatory networks and RNA-seq data, for the first time, [5], [6], [7], [8]. This adaptation represents a significant innovation in terms of a mathematical structure gaining biological meaning, [3], [4], [5], [6].
- **Relating the Precision Matrix and Conditional Dependencies**: The study mathematically formulates that the precision matrices (Σ^{-1}) of gene expression covariance matrices represent conditional dependencies in gene regulatory networks, [5], [7], [8], [10]. This approach, going beyond traditional correlation analyses, allows for the modeling of direct interactions between genes (with the effect of all other genes removed), [4], [6], [9], [11].
- **Disentangling Structural Conservation from Parametric Change**: For the first time in the literature, the qualitative (network topology) and quantitative (expression levels) components of the difference between two biological states have been separated using the Hyper G test, [5], [6], [8], [3]. This distinction offers a fundamental innovation in understanding cancer development and evolutionary processes, [4], [5], [6], [7].
- **Multi-Dimensional Interpretation Framework**: The study provides an integrated framework interpreting Hyper G test results across four different dimensions: mathematical, network theoretical, spectral, and biological, [1], [2], [9], [11]. This multi-dimensional approach allows abstract test statistics to gain biological significance, bringing an interdisciplinary perspective, [4], [5], [6], [7].
- **n-Dimensional Generalization and Applicability to Real Data**: The developed theoretical framework, with its structure generalizable to arbitrary n dimensions, is applicable to real RNA-seq datasets (thousands of genes), [4], [5], [6], [8]. This generalization paves the way for the method's use in practical bioinformatics applications, [5], [7], [8], [10].
- **A New Bridge in the Genotype-Phenotype Relationship**: Through the relationship established between the cell morphologies in Figure 1 and Figure 2 and the mathematical test results, the study visualizes the phenotypic counterparts of network changes at the molecular level, [3], [5], [7], [8]. This creates a strong bridge between abstract mathematical concepts and concrete biological observations, [4], [5], [6], [9].

The Integrated Perspective Brought by the Study:

- **From Mathematics to Biology**: An abstract matrix definition ($A^{-T} = D_1 B D_2$) has been transformed into meaningful biological hypotheses by relating it to the G-matrix

in quantitative genetics, gene networks in systems biology, and RNA-seq data in genomics, [3], [5], [6], [7]. This association, particularly through the precision matrix (Σ^{-1}), has provided a concrete basis for interpreting conditional dependencies, [4], [5], [6], [8]. The cell morphologies in Figure 1 and Figure 2 visualize the phenotypic counterparts of these mathematical distinctions, bridging abstract concepts and concrete biological observations, [1], [2], [9], [11].

- **Distinguishing Structural and Parametric Change**: The test can make this distinction, which is critical for understanding the nature of evolutionary or cancerous transformations, [3], [5], [7], [8]. Using the same network while only gene activation levels change (Scenario A) versus the network being completely rewired (Scenario B) point to very different biological processes, [4], [5], [6], [9]. The variance scaling observed in Scenario A corresponds to situations like an oncogene being overexpressed but its interaction network being conserved; whereas the topological change in Scenario B reflects more complex and aggressive processes such as the silencing of tumor suppressor genes or oncogenes acquiring new targets, [9], [10], [11], [8]. This distinction could be decisive in determining treatment strategies (e.g., choosing between drugs targeting expression levels and those targeting network structure), [3], [5], [6], [7].
- **Opening Doors to New Research Areas**: This framework lays a theoretical foundation for many new research areas, such as cross-species evolutionary architecture comparisons, measuring the stability of gene modules, network-based analyses of drug resistance, and classifying the molecular architecture of tumors in personalized medicine, [1], [2], [4], [5]. Notably, the Hyper-G test, applicable to real RNA-seq data through its n-dimensional generalization, has the potential to be used in identifying subtypes of different cancer types, predicting drug response, and early detection of treatment resistance, [3], [5], [6], [7]. Furthermore, it brings a new approach to cross-species G-matrix comparisons in evolutionary biology, offering quantitative answers to the question of to what extent the genetic correlation structures determining the direction of adaptive evolution are conserved, [4], [6], [8], [9].

In conclusion, Hyper G-matrices and related tests offer a powerful and innovative toolset that combines mathematical rigor with biological interpretation for understanding complex biological systems, [3], [4], [5], [6], [7], [8]. The presented 4×4 simulation serves as a conceptual model for applying these tools to larger, real-world datasets, [8], [9], [10], [11]. This study emphasizes the importance of an interdisciplinary approach, demonstrating how abstract mathematical concepts can provide powerful insights into understanding biological systems, [1], [2], [3], [9], [10], [11]. The developed theoretical calculations and test procedure bring a new perspective to network-based analyses in the field of bioinformatics, offering a unique methodology applicable in many areas, especially cancer research and evolutionary biology, [3], [4], [5], [6], [7], [8]. Future studies aim to apply this test to real cancer datasets and characterize network rewiring patterns in different cancer types, [1], [2], [8], [9], [10], [11].

Acknowledgments:

The authors gratefully acknowledge the editorial office and the anonymous reviewers.

References:

[1] R. A. Horn and C. R. Johnson: *Matrix Analysis*, 2nd edition, Cambridge University Press, New York (2012). https://www.anandinstitute.org/pdf/Roger_A.Horn.%20_Matrix_Analysis_2nd_edition(BookSee.org).pdf

[2] M. Fiedler and F. J. Hall: *G-matrices*, Linear Algebra and its Applications, **436** (3) (2012), 731–741. https://doi.org/10.1016/j.laa.2011.08.001

[3] G. Călugăreanu: *Sylvester matrices*, Special Matrices, **13** (2025). https://doi.org/10.1515/spma-2025-0044

[4] H. Keleş and F. J. Hall: *On Properties Between G-matrices and Hyper G-matrices*, Equations, **5** (2025), 90–99. https://doi.org/10.37394/232021.2025.5.9

[5] H. Keleş and F. J. Hall: *Further Structured Hyper G-matrix Pairs of Matrices*, WSEAS Transactions on Mathematics, **24** (2025), 714–725. https://doi.org/10.37394/23206.2025.24.71

[6] H. Keleş: *On New Developments for Hyper G-Matrices*, Equations, **5** (2025), 23–28. https://doi.org/10.37394/232021.2025.5.4

[7] H. Keleş: *On New Properties of Hyper G-Matrices*, J. Appl. and Pure Math. Vol.

7(2025), No. 5-6, pp. 351–363. https://doi.org/10.23091/japm.2025.351

[8] N. Popitsch, T. Neumann, A. von Haeseler, S. L. Ameres, *Splice-sim: a nucleotide conversion-enabled RNA-seq simulation and evaluation framework*, Genome Biol., **25**, 166 (2024). https://doi.org/10.1186/s13059-024-03313-8

[9] A. Rizvi, P. Camara, E. Kandror, T. Roberts, I. Schieren, T. Maniatis, R. Rabadan, *Single-cell topological RNA-Seq analysis reveals insights into cellular differentiation and development*, Nat. Biotechnol., **35** (2017). https://doi.org/10.1038/nbt.3854

[10] K. R. Kukurba, S. B. Montgomery: *RNA Sequencing and Analysis*, Cold Spring Harb. Protoc., **2015**(11), 951–969 (2015). https://doi.org/10.1101/pdb.top084970

[11] X. Li, S. Zhang, K.-C. Wong: *Single-cell RNA-seq interpretations using evolutionary multiobjective ensemble pruning*, Bioinformatics, **35**(16), 2809–2817 (2018). https://doi.org/10.1093/bioinformatics/bty1056