OXFORD

ONDSA: a testing framework based on Gaussian graphical models for differential and similarity analysis of multiple omics networks

Jiachen Chen [ID][1], Joanne M. Murabito [ID][2,3], Kathryn L. Lunetta [ID][1,*]

[1]Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Crosstown, 3rd floor, Boston, MA 02218, United States
[2]Framingham Heart Study, National Heart, Lung, and Blood Institute and Boston University Chobanian & Avedisian School of Medicine and Boston Medical Center, 73 Mount Wayte Avenue, Framingham, MA 01702, United States
[3]Department of Medicine, Section of General Internal Medicine, Boston University Chobanian & Avedisian School of Medicine and Boston Medical Center, 72 E Concord St, Suite L-516, Boston, MA 02118, United States

*Corresponding author: Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Crosstown, 3rd floor, Boston, MA 02218, United States. E-mail: klunetta@bu.edu

Abstract

The Gaussian graphical model (GGM) is a statistical network approach that represents conditional dependencies among components, enabling a comprehensive exploration of disease mechanisms using high-throughput multi-omics data. Analyzing differential and similar structures in biological networks across multiple clinical conditions can reveal significant biological pathways and interactions associated with disease onset and progression. However, most existing methods for estimating group differences in sparse GGMs only apply to comparisons between two groups, and the challenging problem of multiple testing across multiple GGMs persists. This limitation hinders the ability to uncover complex biological insights that arise from comparing multiple conditions simultaneously. To address these challenges, we propose the Omics Networks Differential and Similarity Analysis (ONDSA) framework, specifically designed for continuous omics data. ONDSA tests for structural differences and similarities across multiple groups, effectively controlling the false discovery rate (FDR) at a desired level. Our approach focuses on entry-wise comparisons of precision matrices across groups, introducing two test statistics to sequentially estimate structural differences and similarities while adjusting for correlated effects in FDR control procedures. We show via comprehensive simulations that ONDSA outperforms existing methods under a range of graph structures and is a valuable tool for joint comparisons of multiple GGMs. We also illustrate our method through the detection of neuroinflammatory pathways in a multi-omics dataset from the Framingham Heart Study Offspring cohort, involving three apolipoprotein E genotype groups. It highlights ONDSA's ability to provide a more holistic view of biological interactions and disease mechanisms through multi-omics data integration.

Keywords: multi-omics networks; differential and similar structures; Gaussian graphical model; false discovery rate

Introduction

The rapid advancement of high-throughput omics technologies has led to the generation of complex, high-dimensional multi-omics datasets including genomics, transcriptomics, proteomics, metabolomics, epigenomics, etc. Graphical models are powerful tools for constructing biological networks from omics data, where the nodes can be genes, proteins, and/or metabolites, and the edges reflect relationships such as co-expression and interactions within or between these omics data [1]. Assuming data with multivariate normality, the Gaussian graphical model (GGM) is a statistical approach for studying dependency networks. In GGMs, two nodes are connected with an edge if and only if they are conditionally dependent given all the other nodes in the network [2]. Unlike correlation networks, which can be quite dense and uninterpretable, GGMs tend to be more sparse since the edges represent direct, rather than indirect, associations [3]. There is

an ever-growing body of literature on GGM estimation in various settings, including omics analyses [1, 3–19].

Differential network analysis is relevant for biomarker discovery and investigation of molecular mechanisms associated with disease onset and progression [20]. Differences in gene, protein, and metabolite networks among two or more clinical conditions may reflect meaningful biological pathways or interactions [20]. We would expect multiple GGMs to be similar when estimated from different conditions of the same type of tissue but also to have some differential structures regulated by distinct but related conditions [21]. In this paper, we focus on identifying biological interactions related to distinct conditions through entry-wise comparisons of multiple precision matrices that represent GGMs. Loosely speaking, recent approaches accounting for similar and differential structures in GGMs depend on three perspectives: joint estimation of multiple GGMs, direct estimation

of differences between GGMs, and separate estimation of multiple GGMs [22]. The spirit of joint estimation of multiple GGMs relies on combining information from all conditions/groups for more accurate estimation within each condition. As a generalization of the well-known graphical lasso method [6], Guo *et al.* [23] and Danaher *et al.* [21] proposed hierarchical penalties with the group lasso or fused lasso for the negative log-likelihood function, encouraging similar coefficients or similar sparsity patterns for precision matrices across groups. Zhang and Wang [24] studied the regression-based neighborhood selection method using the fused lasso penalty. Ha, Baladandayuthapani, and Do [25] developed a pathway-based differential network analysis in genomics that conducted joint estimations by decomposing into global and group-specific components. Cai *et al.* [26] developed a method for joint estimation of multiple precision matrices based on a weighted constrained $L\infty/L1$ minimization. When the focus is on the differences between networks, Zhao, Cai, and Li [27] proposed to estimate the matrix equal to the difference between two precision matrices directly, utilizing the constrained L1 minimization approach. Haslbeck [28] introduced a regression-based method to directly estimate group differences based on moderation analysis via group-specific interaction terms.

Most existing methods for estimating group differences in sparse GGMs only apply to comparisons between two groups, and the challenging problem of multiple testing related to multiple GGMs in terms of structural differences and similarities persists. Even though methods exist for comparisons across more than two groups [21, 26, 28], theoretical justifications for the performance of the estimators have not yet been derived to establish rigorous statistical tests. Furthermore, none of these methods provides false discovery rate (FDR) control in the presence of differential and similar structures regarding more than two groups. These limitations hinder the discovery of significant structural differences and similarities in biological networks across multiple clinical conditions, which is crucial for understanding disease mechanisms and identifying potential biomarkers associated with disease onset and progression. To address these critical gaps, the motivation of this work is to develop a rigorous statistical framework capable of performing simultaneous comparisons across multiple GGMs with proper FDR control.

FDR, introduced by Benjamini and Hochberg [29] for multiple testing problems, is particularly useful in detecting numerous significant comparisons while minimizing the occurrence of false positives. Most approaches for joint estimation of multiple GGMs or direct estimation of differences among them lack theoretical performance guarantees. Therefore, to establish formal statistical tests with FDR control, it is straightforward to consider single GGM estimators incurring asymptotically normal and efficient properties for multiple nonoverlapping groups separately. Liu [15] developed a regression-based single GGM estimator and proposed a procedure for FDR control to simultaneously test the presence of all pairs of nodes in one graph. Xia, Cai, and Cai [30] introduced a multiple testing procedure for differential edges between two groups with FDR control. He *et al.* [31] tested differential network structures between two groups with FDR control based on the well-established estimator of Ren *et al.* [16]. However, jointly comparing more than two sparse GGMs with FDR control remains a complicated problem.

We propose a new two-step statistical testing framework with FDR control for testing both structural differences and similarities across multiple GGMs, with each single GGM estimated by Ren *et al.* [16]. With the implementation and effective application to multi-omics data, we refer to this framework as Omics Networks Differential and Similarity Analysis (ONDSA). ONDSA conducts large-scale multiple testing simultaneously and efficiently, providing a valuable tool for joint comparisons of multiple biological networks that can be leveraged for multi-omics data integration and pathway identification. Unlike many existing methods, ONDSA rigorously performs statistical comparisons across multiple conditions, addressing the methodological gap in multi-group GGM analysis with proper FDR control. Liu's [32] hierarchical method to estimate differential and similar structures under FDR control, which requires procedures with tuning parameters, is a precursor to our method. However, our work focuses on the entry-wise equality of precision matrices (partial covariances), a methodological approach also adopted by many existing differential network analysis studies [21, 27, 30, 31, 33]). This focus on partial covariances continues to be a significant area of research, as evidenced by ongoing studies and advancements [17, 18, 34]. Moreover, we propose distinct test statistics based on a computationally efficient and tuning-free inference method that does not require data-driven approaches to determine tuning parameters, thereby reducing complexity and computational costs [35]. We illustrate the advantages of ONDSA through comprehensive simulations and an application to multi-omics data from the Framingham Heart Study (FHS) Offspring cohort for identifying potential pathways related to neuroinflammation. We conclude with a discussion on how the framework of ONDSA is general and can be extended in both application and methodological perspectives. This well-documented and accessible tool enables researchers to accurately identify both differential and similar omics network structures across multiple conditions, leading to deeper insights into the underlying biological mechanisms and facilitating the discovery of novel biomarkers and therapeutic targets.

Materials and methods
Overview of ONDSA

Here, we provide an overview of ONDSA, reserving the technical details for a later section. The aim of ONDSA is to conduct a two-step rigorous statistical test for identifying differential and similar network structures across precision matrices for multiple groups, with control of FDR at desired levels. We begin by introducing the notation and definitions that will be used in the rest of the paper.

In the context of omics networks, the nodes represent biological entities and the networks are modeled by sparse GGMs. Here, we examine K networks, each corresponding to a distinct group represented by a biological condition and observed over a set of p biological variables (nodes). We consider n_k independent and identically distributed (*i.i.d.*) p-dimensional random vectors, denoted as $\left\{X_d^{(k)} \in R^p : d = 1, \ldots, n_k\right\}$, where each vector represents an individual measurement in group k for $k = 1, \ldots, K$. These vectors follow a group-specific multivariate normal distribution $N(\boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)})$, where $\boldsymbol{\mu}^{(k)}$ is the mean vector and $\boldsymbol{\Sigma}^{(k)}$ is the $p \times p$ covariance matrix. Without loss of generality, we assume $\boldsymbol{\mu}^{(k)} = \mathbf{0}_p$. Let $\Omega^{(k)} = (\boldsymbol{\Sigma}^{(k)})^{-1} = (\omega_{ij,k})$ denote the group-specific precision matrix for group k, where $\omega_{ij,k}$ is the (i,j)-th entry $(1 \leq i, j \leq p)$. The goal of $\omega_{ij,k}$ is to model the conditional dependency between variables i and j in group k while controlling for the influence of the other variables in the network. The joint comparisons conducted by ONDSA are based on group-specific precision matrices
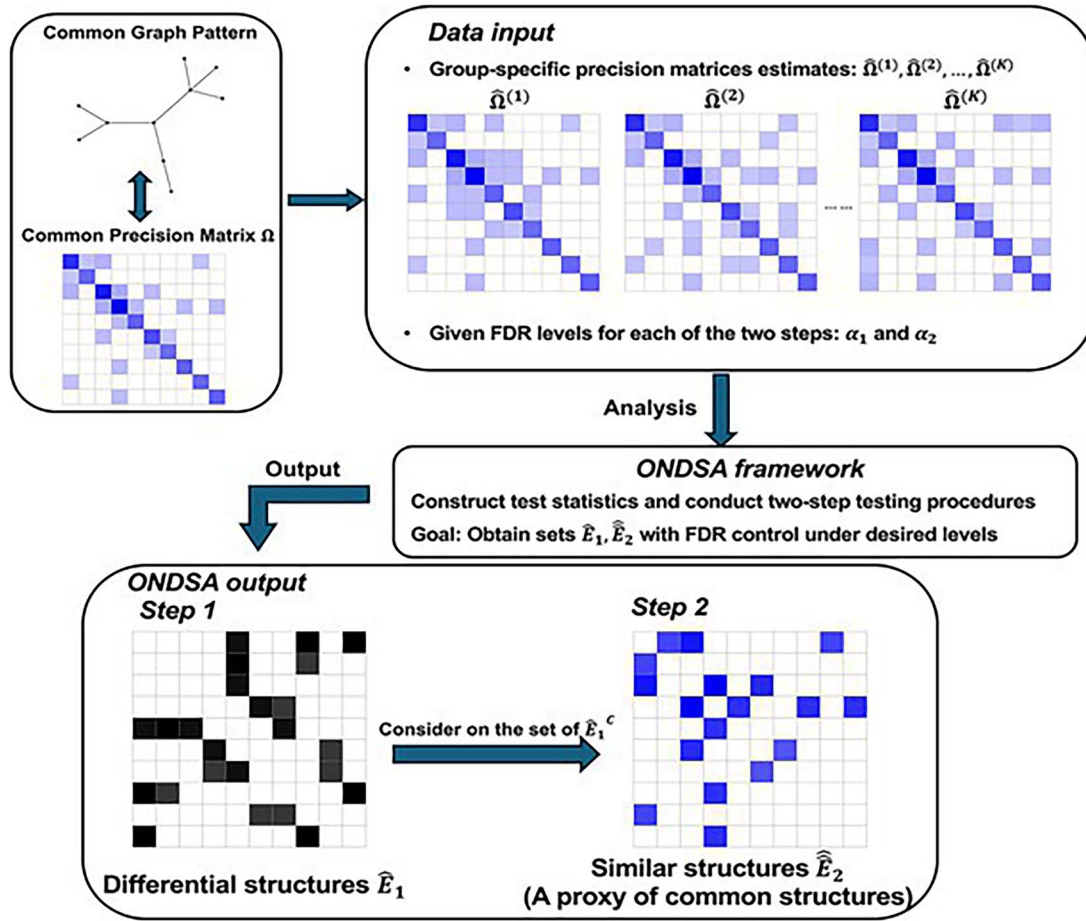
Figure 1. Schematic of the two-step ONDSA.

estimated from each of the K groups using a single GGM inference method [16]. The data harmonization step incorporates all omics and phenotypes of interest into the p nodes, merging them into a dataset with n samples and p variables; here n represents a general notation for the sample size within each single group. The core assumptions are that the data from each group conform to a multivariate normal distribution, and the observations within each group are *i.i.d.* We assume that each group is measured on the same set of biological variables, ensuring that the number of variables p remains the same across all groups. Importantly, we consider high-dimensional settings where the dimension p may grow with the sample size n, accommodating both $(p > n)$ and $(p < n)$ scenarios. The number of groups K is fixed in our analysis. GGMs are useful for modeling omics networks based on continuous variables, assuming that the data can be transformed to approximate normality.

Figure 1 shows the schematic of ONDSA. The estimated group-specific precision matrices $\widehat{\Omega}^{(k)}$ are used as input for ONDSA to describe the network structures for each group. Conducting the two-step testing procedures allows us to obtain the estimated sets for differential and similar structures across multiple nonoverlapping groups, with FDR control at a desired level. The ONDSA R package is available at: https://github.com/jiachenchen322/ONDSA.

Single GGM inference

Anderson [36] showed that the precision matrix can be constructed using linear regression models, and numerous GGM estimation methods have been proposed based on regression coefficients. Sun and Zhang [37] and Ren *et al.* [16] developed a bivariate penalized regression approach focusing on noise level estimation, where an asymptotically efficient estimator for the precision matrix entries ω_{ij} was proposed under sparseness assumptions [16]. To illustrate the estimation for a single group with n samples and p variables, we assume that X is the $n \times p$ omics dataset for a specific biological condition. We are interested in the index set $E = (i, j)$ with $i < j$, and the corresponding entry ω_{ij} of the precision matrix Ω. The conditional Gaussian distribution of the two selected variables given the remaining variables is given by $X_E | X_{E^c} \sim N(-\Omega_{E,E}^{-1}\Omega_{E,E^c}X_{E^c}, \Omega_{E,E}^{-1})$, where $\Omega_{E,E} = \begin{pmatrix} \omega_{ii} & \omega_{ij} \\ \omega_{ij} & \omega_{jj} \end{pmatrix}$. Here, X_E represents the $n \times 2$ matrix corresponding to the index set E, and X_{E^c} represents the $n \times (p - 2)$ matrix excluding the index set E. $\Omega_{C,D}$ represents the submatrix of Ω with row indices C and column indices D. If we regress X_E on X_{E^c}, we will get the equation $X_E = X_{E^c}\beta + \epsilon_E$, where $\beta = \beta_{E^c,E} = (\beta_{E^c,i}, \beta_{E^c,j}) = -\Omega_{E^c,E}\Omega_{E,E}^{-1}$. Here, $\beta = \beta_{E^c,E}$ represents a $(p - 2) \times 2$ coefficient matrix. An approach using scaled lasso that is tuning-free provides the estimate $\widehat{\beta}$ of β [16]. The residual $\widehat{\epsilon}_E = X_E - X_{E^c}\widehat{\beta}$ can be used to obtain an estimator of $\Omega_{E,E}$: $\widehat{\Omega}_{E,E} = \begin{pmatrix} \widehat{\omega}_{ii} & \widehat{\omega}_{ij} \\ \widehat{\omega}_{ij} & \widehat{\omega}_{jj} \end{pmatrix} = (\frac{1}{n}\widehat{\epsilon}_E'\widehat{\epsilon}_E)^{-1}$. The estimator $\widehat{\omega}_{ij}$ is asymptotically efficient within a class of parameter spaces modeling sparse precision matrices:

$$\sqrt{n\widehat{F}_{ij}}(\widehat{\omega}_{ij} - \omega_{ij}) \xrightarrow{D} N(0, 1), \widehat{F}_{ij}/F_{ij} \rightarrow 1. \tag{1}$$

Here, $\widehat{F}_{ij} = (\widehat{\omega}_{ii}\widehat{\omega}_{jj} + \widehat{\omega}_{ij}^2)^{-1}$, and $F_{ij} = (\omega_{ii}\omega_{jj} + \omega_{ij}^2)^{-1}$ represents the Fisher information for estimating ω_{ij}.

This estimation process is crucial for capturing the intricate dependency structures in omics networks. However, there is a trade-off between asymptotic efficiency and sparsity. While the estimators are statistically efficient, they may lead to dense networks, making it more challenging to compare biological relationships across multiple groups.

ONDSA: the proposed two-step framework

The previous section described the GGM estimator for a single group. For nonoverlapping samples from multiple groups, the precision matrix for each group is estimated as $\widehat{\Omega}^{(k)} = (\widehat{\Sigma}^{(k)})^{-1} = (\widehat{\omega}_{ij,k})$ with $K \geq 2$. Below, we provide the details of ONDSA for testing differential and similar structures sequentially.

Step 1: test for differential omics network structures

For omics networks, it is crucial to understand whether specific biological relationships change under different conditions. The multiple testing problem for detecting differential omics network structures across groups is formulated as

$$H_{0ij} : \omega_{ij,1} = \dots = \omega_{ij,K} \text{ vs } H_{1ij} : \text{at least one } \omega_{ij,k} \text{ differs} \quad (2)$$

for $1 \leq i < j \leq p$ and $1 \leq k \leq K$, where the null hypothesis indicates no change in the conditional relationship between variables i and j across the groups.

As shown in (1), under certain conditions, for $1 \leq k \leq K$, $1 \leq i < j \leq p$,

$$\frac{\sqrt{n_k}(\widehat{\omega}_{ij,k} - \omega_{ij,k})}{\sqrt{\widehat{\omega}_{ii,k}\widehat{\omega}_{jj,k} + \widehat{\omega}_{ij,k}^2}} \xrightarrow{D} N(0, 1). \quad (3)$$

This allows us to construct a weighted average $\bar{\omega}_{ij} = \frac{\sum_{k=1}^{K} \lambda_{ij,k}\widehat{\omega}_{ij,k}}{\sum_{k=1}^{K} \lambda_{ij,k}}$ across the K groups, where $\lambda_{ij,k} = \frac{1}{\frac{\widehat{\omega}_{ii,k}\widehat{\omega}_{jj,k} + \widehat{\omega}_{ij,k}^2}{n_k}}$. Inspired by Cochran Q test for heterogeneity, we define $D_{ij} = \sum_{k=1}^{K} \lambda_{ij,k}(\widehat{\omega}_{ij,k} - \bar{\omega}_{ij})^2$, where D_{ij} quantifies the extent to which the relationship between two variables changes across different biological conditions. Under the null hypothesis of no difference, the test statistic D_{ij} follows an asymptotic Chi-squared distribution with $K - 1$ degrees of freedom. To account for the multiple comparisons across the network, we apply the FDR control procedure [38] to adjust for the correlation among the $p(p-1)/2$ test statistics. First, D_{ij} is transformed into a z-value as follows:

$$S_{ij,1} = \Phi^{-1}(\Psi_0(D_{ij})), \quad (4)$$

where $\Psi_0(s)$ corresponds to the cumulative distribution function of the Chi-squared distribution with $K-1$ degrees of freedom and Φ is the cumulative distribution function of the standard normal distribution. This transformation ensures that larger values of $S_{ij,1}$ indicate more substantial changes in edge strengths (conditional dependencies) across groups.

To correct for correlations among entry-wise test statistics caused by overlapping nodes in the network, we apply a conditional FDR adjustment. The conditional FDR is computed as follows: $\text{FDR}(s|A) = \text{FDR}_0(s)(1 + A\frac{|s|\varphi(s)}{\sqrt{2}(1-\Phi(s))})$, where $\text{FDR}_0(s)$ is the standard unconditional FDR under the null $N(0, 1)$, and can be expressed as $\text{FDR}_0(s) = \frac{(1-\Phi(s))p(1-p)/2}{\max(1, \sum_{1\leq i<j\leq p} I(S_{ij,1}\geq s))}$ [38]. The factor A is computed as $A = (P_0 - \widehat{P}_0)/Q_0$, where $P_0 = 2\Phi(1) - 1$, $\widehat{P}_0 = $

$\frac{2\sum_{1\leq i<j\leq p} I(|S_{ij,1}|\leq 1)}{p(p-1)}$, and $Q_0 = \sqrt{2}\varphi(1)$, with $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. The selection of the null, P_0, and Q_0 follows the recommendations of Efron [38]. We denote the adjustment factor as $A(s) = 1 + A\frac{|s|\varphi(s)}{\sqrt{2}(1-\Phi(s))}$. Given a pre-specified $0 < \alpha_1 < 1$, the FDR control procedure for estimating differential omics network structures is performed using the threshold $\widehat{s}_1$, define as

$$\widehat{s}_1 = \inf\left\{s \in \mathbf{R} : \frac{A(s)(1 - \Phi(s))p(p-1)/2}{\max(1, \sum_{1\leq i<j\leq p} I(S_{ij,1}\geq s))} \leq \alpha_1\right\},$$
$$\text{equivalent to FDR}(s|A) \leq \alpha_1. \quad (5)$$

We reject the null hypothesis H_{0ij} if $S_{ij,1} \geq \widehat{s}_1$, indicating a significant difference in edge strength between this specific pair of nodes across the groups.

Failure to adjust for the factor $A(\widehat{s}_1)$ can lead to misleading results, as correlations among test statistics may distort the null distribution [38]. Assume that the set E_1 represents the true set of edges with different edge strengths across groups. The aforementioned procedure produces an estimator $\widehat{E}_1 = \{(i, j) : S_{ij,1} \geq \widehat{s}_1, 1 \leq i < j \leq p\}$, identifying omics networks structures where significant differences across groups are detected.

Step 2: test for similar omics network structures

While differential structures reveal how relationships between variables change, similar omics network structures highlight stable relationships that persist across groups. These persistent relationships can provide insights into core biological pathways that remain unchanged despite varying biological conditions.

For all pairs of nodes, the ground truth categorizes all possible edges into three sets: true differential edges, true common nonzero edges, and true common zero edges that are not of interest. E_1 represents the true differential edges, while its complementary set, E_1^c, can be divided into two subsets:

$$E_2 = \{(i, j) : \omega_{ij,1} = \dots = \omega_{ij,K} \neq 0, 1 \leq i < j \leq p\}; \quad (6)$$

$$E_3 = \{(i, j) : \omega_{ij,1} = \dots = \omega_{ij,K} = 0, 1 \leq i < j \leq p\} \quad (7)$$

where E_2 is the set of true common nonzero edges, and E_3 is the set of true common zero edges. However, due to the limited sample size and theoretical results of minimax rate [16], it is challenging to accurately identify edges belonging to E_2, where the true nonzero $\omega_{ij,k}$ are exactly the same across all groups. In practice, even true common edges in E_2 may not yield identical estimates across groups. Therefore, following the notation in Liu [32], we consider a proxy for E_2: $\widehat{E}_2 = \{(i, j) \in \widehat{E}_1^c : (\omega_{ij,1}, \dots, \omega_{ij,K}) \neq 0\}$, which represents nonzero similar edges or structures that allow for nearly equal $\omega_{ij,k}$ across groups. The estimated proxy set for E_2 should include most common nonzero edges so we would rather recover similar structures instead of the common nonzero structures that have exactly the same $\omega_{ij,k}$ across groups. The identification of $\widehat{E}_2$ involves the following multiple testing problem for all pairs $(i, j) \in \widehat{E}_1^c$:

$$H_{0ij} : (\omega_{ij,1}, \dots, \omega_{ij,K}) = 0 \text{ vs } H_{1ij} : (\omega_{ij,1}, \dots, \omega_{ij,K}) \neq 0, \quad (8)$$

where the alternative hypothesis indicates that the variables i and j share a similar conditional dependency/edge strength across groups.

As defined in Step 1, the weighted average $\overline{\omega}_{ij} = \frac{\sum_{k=1}^{K} \lambda_{ij,k} \widehat{\omega}_{ij,k}}{\sum_{k=1}^{K} \lambda_{ij,k}}$ is computed, where $\lambda_{ij,k} = \frac{1}{\frac{\widehat{\omega}_{ii,k}\widehat{\omega}_{jj,k}+\widehat{\omega}_{ij,k}^2}{n_k}}$ for all pairs $(i,j) \in \widehat{E}_1^c$. Under the null hypothesis, this statistic follows an asymptotic normal distribution $N(0, \frac{1}{\sum_{k=1}^{K} \frac{1}{\frac{\widehat{\omega}_{ii,k}\widehat{\omega}_{jj,k}}{n_k}}})$. The z-value for the test is computed as

$$S_{ij,2} = \Phi^{-1}\left(2\Phi\left(\left| \left(\sum_{k=1}^{K} \frac{1}{\frac{\widehat{\omega}_{ii,k}\widehat{\omega}_{jj,k}}{n_k}} \right)^{\frac{1}{2}} \bar{\omega}_{ij} \right| \right) - 1 \right). \quad (9)$$

Similarly, to correct for correlations between test statistics caused by overlapping network nodes, we apply a conditional FDR adjustment, where $\widehat{P}_0' = |\widehat{E}_1^c|^{-1}\sum_{(i,j)\in\widehat{E}_1^c} I(|S_{ij,2}| \leq 1)$, $A' = (P_0 - \widehat{P}_0')/Q_0$, and $A'(s) = 1 + A'\frac{|s|\varphi(s)}{\sqrt{2}(1-\Phi(s))}$ [38]. Here, $|\widehat{E}_1^c|$ is the number of elements in this set. The FDR control procedure for identifying similar omics network structures across groups is given by the following threshold for a pre-specified $0 < \alpha_2 < 1$:

$$\widehat{s}_2 = \inf\left\{ s \in \mathbf{R} : \frac{A'(s)(1-\Phi(s))|\widehat{E}_1^c|}{\max(1, \sum_{(i,j)\in\widehat{E}_1^c} I(S_{ij,2} \geq s))} \leq \alpha_2 \right\},$$

$$\text{equivalent to } \mathrm{FDR}(s|A') \leq \alpha_2. \quad (10)$$

We reject the null hypothesis $H_{0ij} : (\omega_{ij,1}, ..., \omega_{ij,K}) = 0$ if $S_{ij,2} \geq \widehat{s}_2$, indicating that the edge between this specific pair of nodes shows a significant similar strength across groups. The estimator $\widehat{\widehat{E}}_2 = \{(i,j) : S_{ij,2} \geq \widehat{\widehat{s}}_2, (i,j) \in \widehat{E}_1^c, 1 \leq i < j \leq p\}$ for similar omics network structures is obtained.

Remarks on ONDSA

The procedures in Equations (5) and (10) can be evaluated using the empirical false discovery proportion (eFDP) and empirical FDR (eFDR), which provide insight into the performance of the tests for differential and similar omics network structures. For testing differential structures, the eFDP and eFDR are defined as follows:

$$\mathrm{eFDP1} = \frac{\sum_{(i,j)\in\widehat{E}_1} I\{\omega_{ij,1} = ... = \omega_{ij,K}\}}{\max\left(1, \left|\widehat{E}_1\right|\right)}$$

$$\text{and} \quad \mathrm{eFDR1} = \mathrm{E}\left[\mathrm{eFDP1}\right] \quad (11)$$

For testing similar structures, the corresponding eFDP and eFDR are defined as follows:

$$\mathrm{eFDP2} = \frac{\sum_{(i,j)\in\widehat{\widehat{E}}_2} I\left\{(\omega_{ij,1}, ..., \omega_{ij,K}) = 0\right\}}{\max\left(1, \left|\widehat{\widehat{E}}_2\right|\right)}$$

$$\text{and} \quad \mathrm{eFDR2} = \mathrm{E}\left[\mathrm{eFDP2}\right]. \quad (12)$$

These metrics quantify the control of false discoveries when identifying differential and similar edges in omics networks. Additionally, the empirical power (ePower) measures the ability to detect true differential and similar structures, defined as follows:

$$\mathrm{ePower1} = \mathrm{E}\left(\frac{\sum_{(i,j)\in E_1} I\left\{(i,j)\in\widehat{E}_1\right\}}{|E_1|} \right), \quad (13)$$

$$\mathrm{ePower2} = \mathrm{E}\left(\frac{\sum_{(i,j)\in E_2} I\left\{(i,j)\in\widehat{\widehat{E}}_2\right\}}{|E_2|} \right). \quad (14)$$

Algorithm 1 ONDSA with FDR Control

Input: Estimated group-specific precision matrices, representing omics networks as defined in Equation (1), and the specified FDR control levels α_1 and α_2

Step 1: test for differential omics network structures
(a) Construct test statistics $S_{ij,1}$ in Equation (4).
(b) Conduct the multiple testing procedure on all node pairs as described in Equation (5) and calculate $\hat{s}_1$.
(c) Obtain the estimated set of differential omics network structures: $\hat{E}_1 = \{(i,j) : S_{ij,1} \geq \hat{s}_1, 1 \leq i < j \leq p\}$.
Step 2: test for similar omics network structures
(a) Construct test statistics $S_{ij,2}$ in Equation (9).
(b) Conduct the multiple testing procedure on the set of $\hat{E}_1^c$ as described in Equation (10) and calculate $\hat{s}_2$.
(c) Obtain the estimated set of similar omics network structures: $\hat{E}_2 = \{(i,j) : S_{ij,2} \geq \hat{s}_2, (i,j) \in \hat{E}_1^c, 1 \leq i < j \leq p\}$.

Output: Estimated sets $\hat{E}_1$ and $\hat{E}_2$ for differential and similar omics network structures across groups.

The complete procedure of ONDSA with FDR control is summarized in Algorithm 1. The goal is to estimate the sets E_1 and E_2, corresponding to the differential and similar structures in omics networks across multiple groups, which investigates both varying and consistent biological relationships across conditions. As previously mentioned, when the differences in edge strengths across groups are extremely small and the sample size n_k is limited, detecting such differences becomes challenging [16]. Specifically, true differential edges in E_1 with small differences across groups may not reach statistical significance and may be classified as similar edges in $\widehat{\widehat{E}}_2$. Similarly, even with FDR control, some true zero edges in E_3 may be falsely identified as nonzero edges and misclassified into $\widehat{\widehat{E}}_2$ due to estimation errors and the high dimensionality of the data, especially when sample sizes are limited [16, 32].

Under some circumstances, it may be of interest to compare the K omics networks for entry-wise partial correlations rather than partial covariances/precision matrix entries [21, 32]. In the Supplementary Materials, we further show the extension of ONDSA to work on the scale of partial correlations.

Simulation setup

We demonstrate the performance of the ONDSA framework through simulation studies, evaluate its empirical FDR and power as previously defined, and compare it with two existing methods. The single GGM estimation in ONDSA is straightforward to implement using `FastGGM` [39], an efficient R package developed for the asymptotically normal and efficient estimator established by Ren *et al.* [16]. The scaled-lasso penalty level $\lambda = \sqrt{(2/n)\log(p/\sqrt{n})}$ is provided as the default value with statistical justifications, where p represents the number of variables (nodes) and n is used as a general notation for the sample size within each single group. Our implementation of ONDSA is developed into an R package. Installation, dependencies, and detailed usage instructions for reproducible research are available on GitHub.

We design simulation scenarios to assess the proposed ONDSA framework under various conditions, using network structures commonly applied in GGM simulations [5, 10, 27, 32, 40]. We consider both low-dimensional ($n \geq p$) and high-dimensional ($n < p$) settings with $K = 3$ groups. Simulated data sets for

these three groups are generated from multivariate normal distributions with different network structures, including various common graphs and individual edge structures, under balanced and imbalanced group scenarios. The common graph precision matrix is denoted by Ω. The corresponding common adjacency matrix is denoted by $\Gamma = (\gamma_{ij})$, where each entry is binary, representing the presence (1) or absence (0) of an edge between nodes i and j. Group-specific precision matrices are denoted by Ω_k. The simulation pipeline consists of three steps. First, a common graph structure is generated for all groups using random, hub, band, or scale-free graphs. Second, individual group-specific edges are assigned to each group by randomly selecting either individual edges or block structures. Group-specific precision matrices Ω_k are obtained from these two steps. Finally, we invert these precision matrices to covariance matrices and generate n_k samples from a multivariate normal distribution $N(0_p, \Omega_k^{-1})$ for $k = 1, 2, 3$. The common graph and individual edge structures used to construct the symmetric precision matrices are defined as follows.

(1) Common graph structures

(a) Random graph: Each pair of nodes is randomly selected as an edge denoted as $\gamma_{ij} = \gamma_{ji} = 1$ with probability 0.02 when $n \geq p$ or 0.008 when $n < p$, and $\gamma_{ij} = \gamma_{ji} = 0$ otherwise.
(b) Hub graph: The p nodes are evenly divided into $max(10, p/20)$ disjoint clusters. In each cluster, there is a central node i, and for all other nodes j within the same cluster, $\gamma_{ij} = \gamma_{ji} = 1$ when $j \neq i$, and $\gamma_{ij} = \gamma_{ji} = 0$ otherwise.
(c) Band graph: For all pairs of nodes, $\gamma_{ij} = \gamma_{ji} = 1$ if $1 \leq |i - j| \leq 2$, and $\gamma_{ij} = \gamma_{ji} = 0$ otherwise.
(d) Scale-free graph: This graph is generated by the Barabási–Albert algorithm. Based on the initial two connected nodes, each new node is connected to one existing node, with the probability of connection being proportional to the degree of the existing node.

Then we set the diagonal entries of Γ to 0 and the smallest eigenvalue of $\Gamma \times 0.3$ is denoted as e. The positive definite common precision matrix is obtained as $\Omega = \Gamma \times 0.3 + (|e| + 0.1 + 0.1)I_p$, which is implemented by the R package huge [41] (v.1.3.5). The strength of common edges is set to 0.3.

(2) Individual edge structures

(a) Random: For each group k, the zero entries in the common graph Ω are randomly selected as group-specific individual edges with a probability of $200/p^2$. The nonzero entries $\omega_{ij,k}$ are drawn from a uniform distribution, either $U(0.2, 0.4)$ or $U(-0.4, -0.2)$ with equal probability of 0.5.
(b) Block-wise: Group-specific individual edges are introduced by adding a block diagonal matrix to the common graph precision matrix. Each block is a 4×4 symmetric matrix, and is repeated $p/4$ times along the diagonal. For the off-diagonal entries within each block, the probability of being nonzero is 0.4 for each pair of nodes, and the corresponding $\omega_{ij,k}$ is generated from a uniform distribution $U(0.2, 0.4)$.

The average absolute strength of the individual edges is 0.3. If, in any instance, the group-specific precision matrix does not satisfy the condition of positive definiteness, it is regenerated until the condition is satisfied.

Compared methods

As discussed in the Introduction, few existing methods allow for the joint comparison of more than two sparse GGMs. Among these, the joint graphical lasso (JGL) [21] and the moderated network model (MNM) [28] are well implemented for applications. Our proposed FDR control framework, ONDSA, tests for differential and similar edges in precision matrices for omics networks within a single analysis. This approach is comparable with the joint estimation methods JGL and MNM in that joint comparisons could be conducted effectively. Therefore, simulations are conducted to compare the performance of the ONDSA framework with these two alternatives. JGL is implemented by the R-package JGL (v.2.3.1), and MNM is implemented by the R-package mgm (v.1.2–13).

Multi-omics dataset from FHS Offspring cohort

To demonstrate the application of the ONDSA framework, we use data from the FHS Offspring cohort [42] at Exam 7 to investigate associations within multi-omics networks, focusing on two omics data: circulating immune cell abundance and plasma (blood) protein biomarkers. Peripheral blood mononuclear cell samples from 608 dementia-free individuals were profiled for 43 primary immune cell phenotypes [43], and plasma samples from the same individuals were profiled for 68 protein biomarkers using the OLINK Inflammation panel (https://olink.com/). For illustrative use, we ignore the kinship relationships among the sample individuals.

The apolipoprotein E (*APOE*) gene remains the strongest genetic risk factor for Alzheimer's disease (AD) [44]. The *APOE* $\epsilon4$ allele significantly increases AD risk, while the $\epsilon2$ allele has protective effects compared with the common $\epsilon3$ allele. Previous studies, such as Soares *et al.* [45], demonstrated variations in plasma biomarker profiles among individuals with different *APOE* genotype carrier groups in AD dementia patients and cognitively normal participants. Immune activation and fluctuations in circulating immune cells have been linked to the production of cytokines and chemokines involved in neuroinflammation [46].

We are particularly interested in analyzing the multi-omics networks to identify conditionally dependent pairs between inflammatory protein biomarkers and immune cell phenotypes that are differential or similar across the three *APOE* genotype groups, with implications for understanding the mechanisms underlying cognitive aging. We exclude individuals with *APOE* $\epsilon2\epsilon4$ genotype or those with missing values and classified the remaining participants into three groups: $\epsilon2$ carriers ($\epsilon2\epsilon2$ and $\epsilon2\epsilon3$ genotypes, n=78), $\epsilon4$ carriers ($\epsilon3\epsilon4$ and $\epsilon4\epsilon4$ genotypes, n=121), and $\epsilon3\epsilon3$ genotypes (n=409). Age and sex information were ascertained at Exam 7. Cytomegalovirus (CMV), a common pathogen known to affect T cell subsets in infected individuals, was also measured. Before performing GGM estimations, we adjust the 111 nodes, including immune cell phenotypes and proteins, for sex, age, and CMV level effects using linear regression. Then, we use the ONDSA framework to estimate differential and similar omics network structures across the three *APOE* groups under FDR control levels $\alpha_1 = \alpha_2 = 0.05$.

Results
Simulation studies
Performance of ONDSA for balanced groups

For each scenario, we generate a balanced group design with $n_1 = n_2 = 333, n_3 = 334$ independent samples from group-specific precision matrices with $p = 100, 500$ for 100 replicates, respectively. The empirical FDR and power for testing differential omics network structures are denoted as eFDR1 and ePower1, while those for testing similar omics network structures are denoted

as eFDR2 and ePower2 (see Equations (11–14)). We obtain the sets $\widehat{E}_1$ and $\widehat{E}_2$ for each replication, and compute **eFDR1**, **eFDR2**, **ePower1**, and **ePower2** using the corresponding equations. We plot eFDR1 and eFDR2 against a reference line $f(\alpha) = \alpha$ to assess if the proposed ONDSA framework effectively controls the FDR at the desired level. We consider FDR control at various levels, with $\alpha = \alpha_1 = \alpha_2 = 0.05, 0.1, 0.15, 0.2$. The empirical FDR and power curves are shown with α value as the x-axis.

Empirical results for all eight scenarios, combining the four common omics network graph structures and two individual edge structures, are presented in Figs 2 and 3 and Figs S1 and S2. In low-dimensional settings ($p = 100$, Figs 2 and 3), our proposed ONDSA effectively controls the FDR at both steps and achieves especially conservative performance when the true common omics network graph structures are random, hub, and scale-free graphs (Fig. 2). Slight inflation in eFDR2 occurs when identifying similar structures in the band graph, likely due to the violation of sparseness conditions of common edges in the omics networks (Fig. 2(c)(g)). The ePower2 for testing similar structures is close to 1 across all scenarios, while the ePower1 for testing differential structures ranges from ˜0.25 to ˜0.9, depending on α levels and scenarios (Fig. 3). The lower ePower1 can be attributed to several differential edges whose differences across groups are near 0. Therefore, they may be misclassified as similar edges, and fewer true positives are retained. Band and hub graphs with block-wise individual edges are particularly susceptible to this issue due to the specific graph structures (Fig. 3(b)(c)(f)(g)). In high-dimensional settings ($p = 500$, Figs S1 and S2), eFDRs are slightly inflated, particularly for eFDR2, when identifying similar omics network structures due to finite sample estimation bias (Fig. S1). Scenarios with scale-free common network structures are detected accurately, with well-controlled eFDRs at both steps (Fig. S1(d)(h)). The ePower2 for testing similar structures may show modest overestimation due to slight inflation in eFDR2, though it still maintains a generally high level. The ePower1 for testing differential structures may be lower than 0.5 when the α is controlled at a lower level, which is also explained by the limited sample size when $n < p$ (Fig. S2).

Performance of ONDSA for imbalanced groups

For each scenario, we generate three groups of independent samples with sizes n_1, n_2, and n_3 in an imbalanced design with $p = 100, 500$ for 100 replicates. The total sample size remains 1000, while the group sizes vary across five imbalanced configurations: $n_1 = 50, n_2 = 150$, and $n_3 = 800$; $n_1 = 100, n_2 = 150$, and $n_3 = 750$; $n_1 = 150, n_2 = 250$, and $n_3 = 600$; $n_1 = 200, n_2 = 300$, and $n_3 = 500$; $n_1 = 300, n_2 = 300$, and $n_3 = 400$. Setting FDR control levels at $\alpha_1 = \alpha_2 = 0.05$, we compute eFDR1, eFDR2, ePower1, and ePower2 as defined earlier, and assess how these metrics change as the group design becomes less imbalanced. We plot eFDR1, eFDR2, ePower1, and ePower2 versus the group sample sizes as the x-axis for each scenario, respectively, in Figs 4 and S3.

In low-dimensional settings (Figs 4(a)–(d) and S3(a)–(d)), when the three groups are extremely imbalanced ($n_1 = 50, n_2 = 150$, and $n_3 = 800$), the eFDR1 for identifying differential structures is substantially higher than the expected level, ranging from ˜0.2 to ˜0.7. However, eFDR1 decreases drastically when $n_1 = 100$, $n_2 = 150$, and $n_3 = 750$, and is controlled near or below 0.05 as the groups become more balanced. The corresponding ePower1 for differential structures is <0.3 when the smallest group size is around 100 or smaller, but increases to ˜0.5 as the group sizes become more balanced, except in scenarios with a band graph and block-wise individual edges (Fig. S3(c)). This extremely low

ePower1 may be attributed to the violation of the sparseness conditions for this specific graph structure. Furthermore, when individual edges follow block-wise structures (Figs S3(a)–(d)), ePower1 initially decreases slightly before increasing as the smallest group size grows, reflecting the interacting effects between block-wise edge structures and imbalanced group sizes. In contrast, eFDR2 for identifying similar structures remains only slightly inflated in the most imbalanced designs and decreases to or below 0.05 as the group design becomes more balanced. The corresponding ePower2 increases as groups become more balanced and remains at high levels ˜0.9. Therefore, in imbalanced designs where there are groups comprising <15% of the total sample size, particularly those with modest sample sizes, the identified differential network edges should be interpreted with caution, while conclusions regarding similar structures tend to be more reliable.

In high-dimensional settings (Figs 4(e)–(h) and S3(e)–(h)), the imbalanced design shows a moderate to slight inflation in eFDRs for identifying differential (eFDR1) and similar (eFDR2) network structures, especially for eFDR2 when the smallest group has a sample size of 100 or less. However, eFDR1 is generally well controlled around or below 0.05, and ePower1 increases to ˜0.5 as the groups become more balanced. This indicates that more differential edges are being correctly identified, though false positives may still occur due to the modest sample sizes in high-dimensional settings. For identifying similar edges, eFDR2 steadily decreases to around 0.05 in more balanced designs but may remain slightly inflated as sample sizes are limited. The ePower2 remains high from ˜0.8 to ˜0.9 even in imbalanced designs. Different from low-dimensional settings, conclusions on both differential and similar omics network structures should be interpreted cautiously, particularly when dealing with modest group sizes, including imbalanced groups with <15% of the total sample. Although slight inflation in eFDR1 is present, differential edges identified are more likely to be accurate compared with low-dimensional settings. However, ePower1 still suffers from the limited sample sizes.

Comparisons with existing methods

We consider the low-dimensional settings ($n_1 = n_2 = 333, n_3 = 334, p = 100$) and high-dimensional settings ($n_1 = n_2 = n_3 = 150, p = 180$) for the eight scenarios of graph structures, each with 100 replicates. The proposed ONDSA framework is applied with FDR control levels $\alpha_1 = \alpha_2 = 0.05$ and requires no tuning procedures. We apply JGL with fused graphical lasso using tuning parameters from cross-validation on the first replicate, then use them for all subsequent replicates. Similar network structures in JGL are defined by edges with absolute differences in precision matrix entries of <0.01. For MNM, we use the extended Bayesian information criteria for model selection, and transform the regression estimates to precision matrix entries for identifying differential and similar structures. To assess the overall performance of methods, the true positive rate (TPR) and false positive rate (FPR) for recovering the differential and similar structures in omics networks are denoted as TPR1, FPR1, TPR2, and FPR2, respectively. The average TPR and FPR for both steps are presented in Table 1, and the conclusions are further validated by empirical 95% confidence intervals.

As shown in Table 1, our proposed ONDSA framework outperforms both JGL and MNM on TPR2 and FPR2 in identifying similar structures across nearly all scenarios of omics networks. For identifying differential structures, JGL demonstrates a higher TPR1 but its FPR1 is also higher, while ONDSA has a higher or similar TPR1
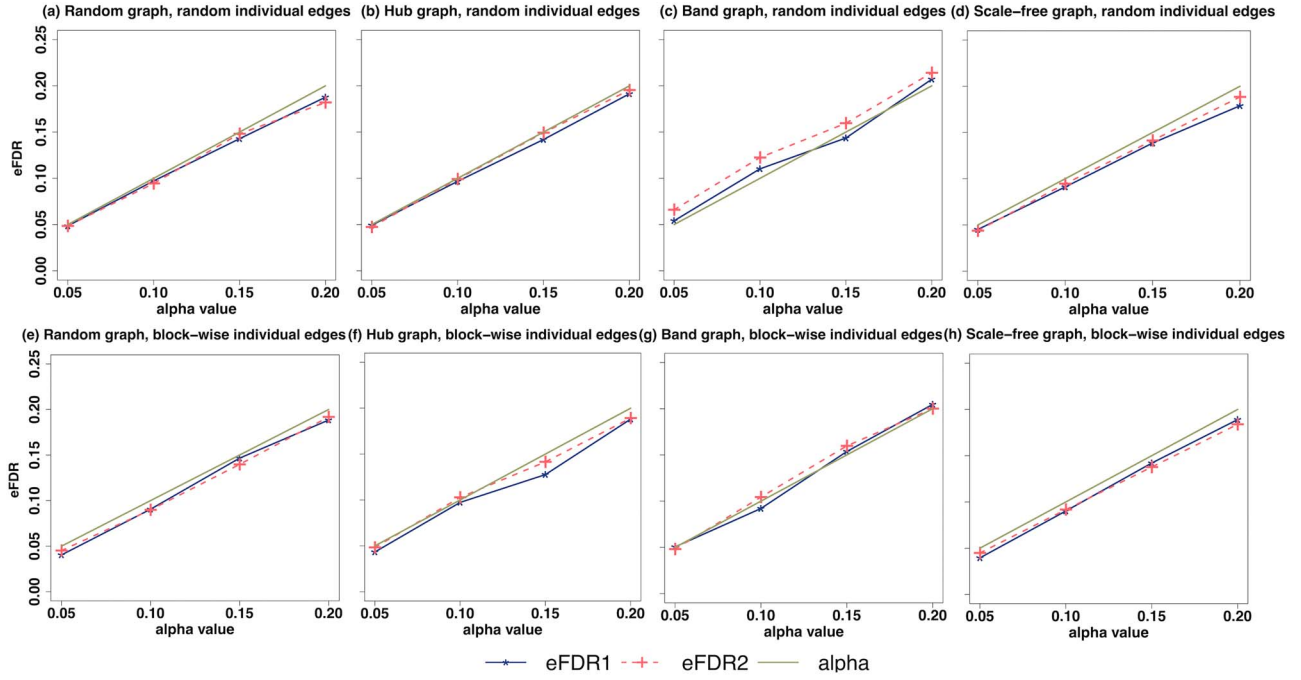
Figure 2. **Empirical FDR performance of the ONDSA framework for p = 100 with balanced group sizes.** We consider eight scenarios of the combination of four common graph structures and two individual edge structures. The x-axis denotes the α value and the y-axis denotes the empirical FDR (eFDR) for testing differential structures (eFDR1) and similar structures (eFDR2). $\text{eFDP1} = \frac{\sum_{(i,j) \in \widehat{E}_1} I\left\{\omega_{ij,1} = \ldots = \omega_{ij,K}\right\}}{\max(1, |\widehat{E}_1|)}$ and $\text{eFDR1} = \text{E}[\text{eFDP1}]$. $\text{eFDP2} = \frac{\sum_{(i,j) \in \widehat{E}_2} I\left\{(\omega_{ij,1}, \ldots, \omega_{ij,K}) = 0\right\}}{\max(1, |\widehat{E}_2|)}$ and $\text{eFDR2} = \text{E}[\text{eFDP2}]$.
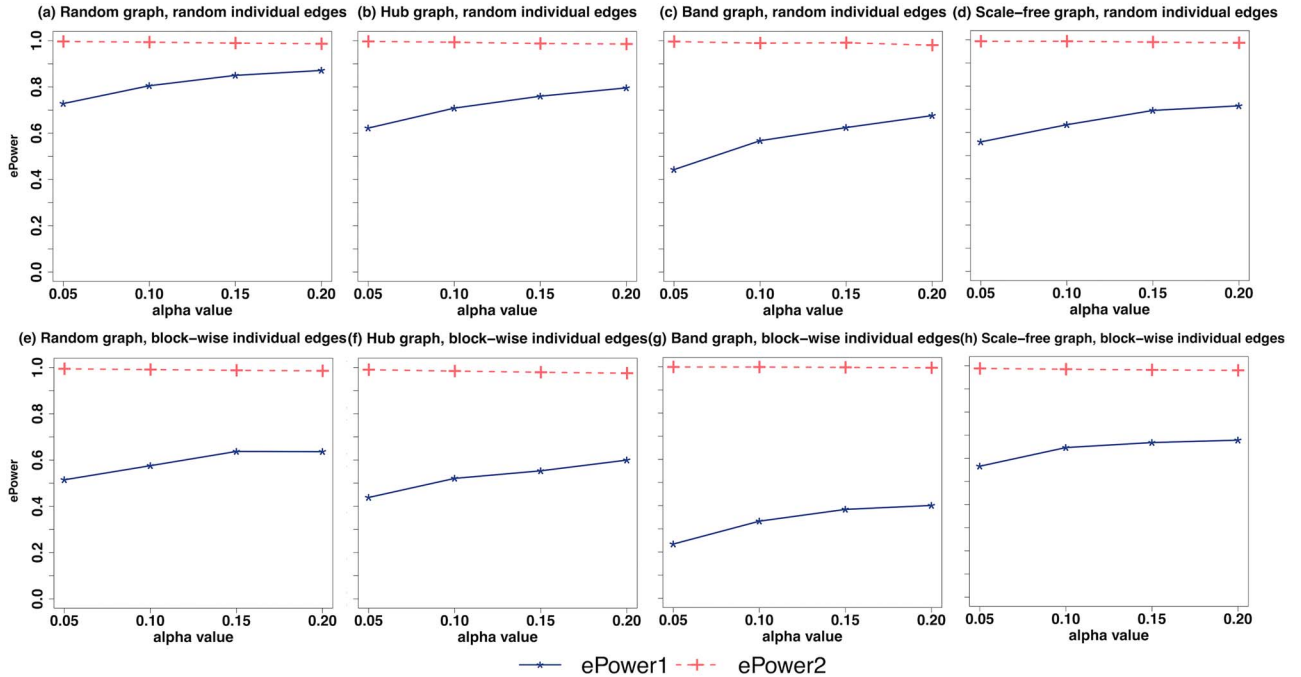

Figure 3. **Empirical power performance of the ONDSA framework for p = 100 with balanced group sizes.** We consider eight scenarios of the combination of four common graph structures and two individual edge structures. The x-axis denotes the α value and the y-axis denotes the empirical power (ePower) for testing differential structures (ePower1) and similar structures (ePower2). $\text{ePower1} = \text{E}\left(\frac{\sum_{(i,j) \in E_1} I\left\{(i,j) \in \widehat{E}_1\right\}}{|E_1|}\right)$, $\text{ePower2} = \text{E}\left(\frac{\sum_{(i,j) \in E_2} I\left\{(i,j) \in \widehat{E}_2\right\}}{|E_2|}\right)$.
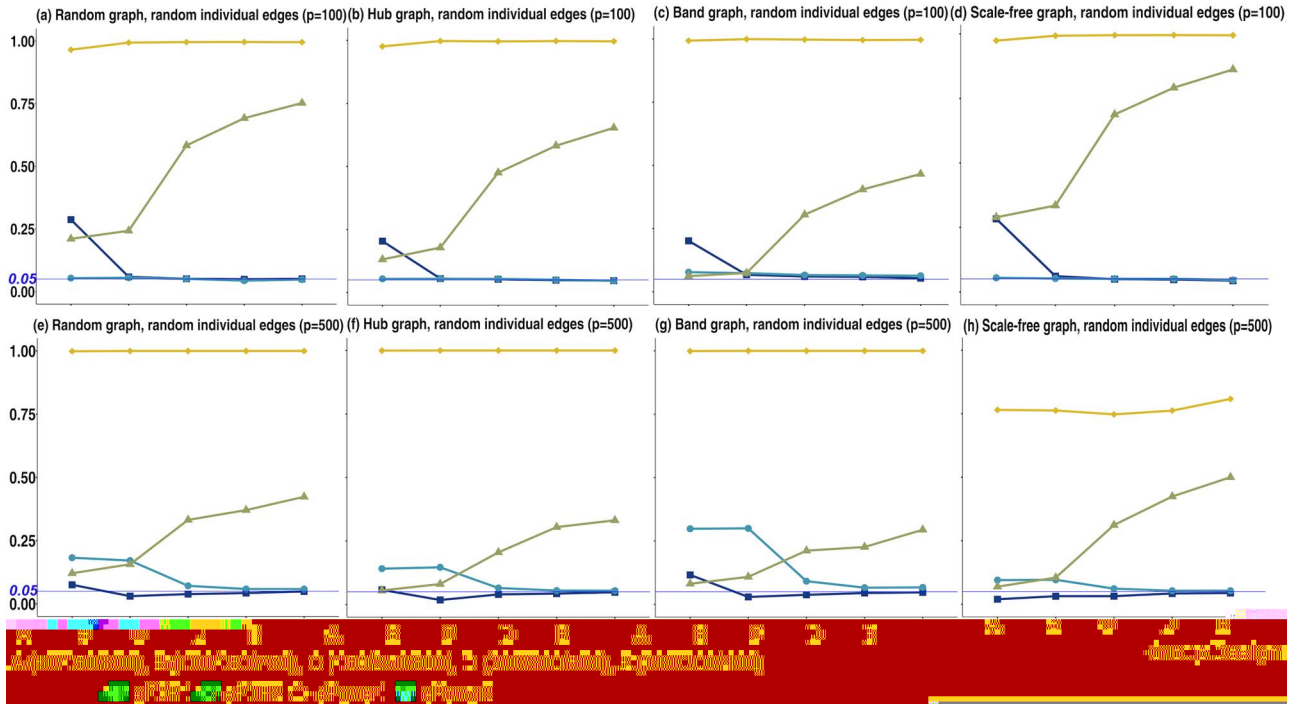
Figure 4. **Performance of the ONDSA framework for p = 100/500, evaluated under imbalanced group sizes and FDR control levels at $\alpha_1 = \alpha_2 = 0.05$, for scenarios with random individual edges.** Each panel shows the empirical FDR (eFDR) and empirical power (ePower) across five imbalanced group size configurations A-E.

with better FPR1 compared with MNM. These trends are consistent across both low-dimensional and high-dimensional settings, and are confirmed by empirical 95% confidence intervals.

Our simulation scenarios provide a comprehensive comparison across methods, evaluating performance under various realistic conditions. We assume different network sizes in both low-dimensional and high-dimensional settings, illustrating the distinct challenges posed by large-scale datasets ($n < p$) versus traditional datasets ($n > p$). In these scenarios, JGL shows weaker performance with higher FPR1 and lower TPR2, especially in high-dimensional settings. Similarly, MNM struggles with lower TPR2 in these settings. In contrast, ONDSA demonstrates consistently superior performance across both network sizes, maintaining robust results regardless of the dimensionality of the data.

We also explore different sparsity levels through various common graph structures. Random and scale-free graphs adhere to the sparsity conditions ideal for ONDSA, as mentioned by [16], while hub and band graphs, being denser, do not meet these sparsity conditions. This is reflected in Table 1, where ONDSA exhibits slightly reduced TPR1 in scenarios with hub or band graphs, highlighting the impact of deviations from ideal sparsity conditions on ONDSA's performance. Despite these challenges, ONDSA still outperforms JGL and MNM in overall effectiveness.

Furthermore, we consider different noise levels through two distinct individual edge structures and various individual edge strength levels. The individual edge structures introduce variations in the group-specific precision matrices, simulating different "noise levels" that diverge from the shared common network graph structure. Random individual edges, which are more scattered, and block-wise individual edges, which tend to form smaller clusters, pose unique challenges. The results in Table 1 indicate that ONDSA shows robust performance for all metrics compared with its alternative, even in scenarios with block-wise individual

edges, which tend to introduce more structured noise. To further evaluate the influence of varying individual edge strengths, we focus on scenarios where the common graph is scale-free and consider cases where the average absolute strength of individual edges is lower than, equal to, or greater than that of the common edges. Detailed descriptions of these additional scenarios are provided in the Supplementary Materials. Table 1 reflects the scenarios where the average absolute strength of individual edges is equal to that of the common edges. As shown in Table S1, when the average absolute strength of individual edges is lower than that of common edges, ONDSA's TPR1 is similar to MNM. However, as the average absolute strength of individual edges increases, TPR1 improves significantly, matching the performance of JGL. Additionally, ONDSA consistently outperforms both JGL and MNM in terms of FPR1, TPR2, and FPR2 while also demonstrating a better balance of FPR1 to TPR1 compared with JGL. These patterns suggest that while ONDSA's TPR1 may not be as high when individual edge strengths are weaker than common edge strengths, its overall competitive performance remains evident across other metrics.

In addition to evaluating performance metrics, we also assess the computational efficiency of each method. For this analysis, we record the computational time required for 100 replicates from each of the eight scenarios within low-dimensional and high-dimensional settings, respectively. We then compute the average computational time required for one replicate, separately for each method—ONDSA, JGL, and MNM—across these eight scenarios. Table 2 provides a straightforward comparison of computational efficiency across methods, underscoring ONDSA's advantages in significantly reducing computational cost.

In conclusion, the proposed ONDSA framework demonstrates excellent performance by maintaining reasonably high TPRs, ranging from ~0.45 to ~0.99, in both discovering differential and

Table 1. Comparison of ONDSA with other existing methods on estimating structural differences and similarities across omics networks

Individual edge structure	Common graph structure	Methods	TPR1	FPR1	TPR2	FPR2
$K = 3, n_1 = n_2 = 333, n_3 = 334, p = 100$						
Random	Random	ONDSA	0.735	0.002	0.997	0.006
		JGL	1.000	0.248	0.281	0.006
		MNM	0.598	0.003	0.910	0.013
	Hub	ONDSA	0.625	0.002	0.998	0.007
		JGL	0.982	0.180	0.373	0.010
		MNM	0.460	0.004	0.877	0.012
	Band	ONDSA	0.446	0.002	0.993	0.012
		JGL	0.995	0.219	0.316	0.010
		MNM	0.465	0.007	0.820	0.013
	Scale-free	ONDSA	0.863	0.002	0.996	0.005
		JGL	0.999	0.258	0.255	0.007
		MNM	0.627	0.004	0.870	0.015
Block-wise	Random	ONDSA	0.798	0.001	0.999	0.004
		JGL	0.993	0.174	0.306	0.005
		MNM	0.673	0.002	0.960	0.005
	Hub	ONDSA	0.738	<0.001	0.998	0.006
		JGL	0.982	0.128	0.207	0.004
		MNM	0.704	0.004	0.835	0.005
	Band	ONDSA	0.423	<0.001	0.999	0.011
		JGL	0.986	0.236	0.243	0.010
		MNM	0.660	0.004	0.846	0.006
	Scale-free	ONDSA	0.843	0.001	0.996	0.003
		JGL	0.986	0.172	0.288	0.005
		MNM	0.691	0.003	0.931	0.005
$K = 3, n_1 = n_2 = n_3 = 150, p = 180$						
Random	Random	ONDSA	0.836	<0.001	0.999	0.001
		JGL	1.000	0.315	0.187	0.006
		MNM	0.689	0.001	0.958	0.001
	Hub	ONDSA	0.454	<0.001	0.997	0.001
		JGL	1.000	0.275	0.241	0.005
		MNM	0.678	<0.001	0.973	<0.001
	Band	ONDSA	0.638	<0.001	0.999	0.003
		JGL	1.000	0.296	0.223	0.007
		MNM	0.669	<0.001	0.947	0.001
	Scale-free	ONDSA	0.881	<0.001	0.668	0.001
		JGL	1.000	0.306	0.175	0.006
		MNM	0.695	<0.001	0.534	0.001
Block-wise	Random	ONDSA	0.824	<0.001	0.997	0.001
		JGL	0.995	0.347	0.182	0.007
		MNM	0.673	0.001	0.962	0.001
	Hub	ONDSA	0.772	<0.001	0.854	0.003
		JGL	0.998	0.447	0.140	0.011
		MNM	0.739	0.001	0.735	0.002
	Band	ONDSA	0.607	<0.001	0.880	0.002
		JGL	0.990	0.413	0.083	0.009
		MNM	0.595	0.002	0.633	0.002
	Scale-free	ONDSA	0.853	0.001	0.705	0.001
		JGL	0.996	0.347	0.174	0.007
		MNM	0.657	<0.001	0.617	0.001

Note: ONDSA, our proposed framework with FDR control; TPR1, true positive rate for identifying differential structures; FPR1, false positive rate for identifying differential structures; TPR2, true positive rate for identifying similar structures; FPR2, false positive rate for identifying similar structures.

similar network structures with negligible FPRs. The proposed ONDSA framework has similar or higher TPRs with similar FPRs compared with MNM, and a better balance of FPR to TPR than JGL under all the simulation scenarios. The robustness to various realistic conditions, combined with its computational efficiency, makes ONDSA a superior option for researchers aiming to analyze complex omics networks efficiently and accurately.

Application of ONDSA to FHS Offspring cohort multi-omics data

The FHS Offspring cohort is well characterized for cognitive aging outcomes. This cohort provides a unique opportunity to explore the mechanisms of neuroinflammation in cognitive aging by leveraging multi-omics data, specifically circulating immune cell abundance and inflammatory protein biomarkers. Our previous

Table 2. Average computational time (in minutes) for one replicate across the eight scenarios considered in Table 1

Settings	ONDSA	JGL	MNM
Low-dimensional settings	0.04	0.24	2.39
High-dimensional settings	0.12	0.70	11.26

Note: ONDSA, our proposed framework with FDR control. Low-dimensional settings: $K=3$, $n_1=n_2=333$, $n_3=334$, $p=100$; high-dimensional settings: $K=3$, $n_1=n_2=n_3=150$, $p=180$.

Table 3. Differential edges across *APOE* genotype groups identified by the ONDSA framework for the illustrative FHS data

Differential edge	$\epsilon2$ carrier	$\epsilon4$ carrier	$\epsilon3\epsilon3$
CD4 Tn - CD4 Tcm	−24.11	−7.97	−16.31
CD8+ T cells[1] - CD4/CD8	118.72	93.77	53.90
CD8 Tn - CD8 Tn/Tm	−40.57	−28.94	−16.64
NK cells - CD38+ T cells	2.22	3.98	0.93
NK cells - CD8+Tc17/Treg	−0.86	3.27	−1.39
CD14 total[2] - NCM	−9.78	−7.89	−4.11
CM - NCM	11.95	10.84	5.81

Note: Positive edge strength indicates negative dependency, while negative edge strength indicates positive dependency. ; Tn, T naïve cells; Tcm, T central memory cells; CD4/CD8, CD4 T helper cells / CD8 Cytotoxic T cells; Tm, memory T cells; NK cells, natural killer cells; Tregs cells, CD8 T regulatory cells. [1] CD8 Cytotoxic T cells. [2] Total monocytes.

work has demonstrated that inflammatory protein biomarkers are associated with cognitive function, as well as the risk of incident dementia and AD [47]. We also showed that the activation of innate and adaptive immunity, along with changes in the composition of circulating immune cells, especially subtypes of immunosenescence, is associated with aging [43]. In this application, we focus on the cross-talk among immune cells and inflammatory proteins, stratified or shared across different *APOE* genotypes, to enhance our understanding of neuroinflammatory mechanisms contributing to cognitive aging.

The three *APOE* genotype groups are $\epsilon2$ carriers (n=78), $\epsilon4$ carriers (n=121), and $\epsilon3\epsilon3$ genotypes (n=409). Using ONDSA to analyze these multi-omics networks, we identify seven differential edges across *APOE* genotype groups (Table 3), all of which are among immune cell pairs, i.e. none of them are between protein and immune cell nodes or between protein pairs. We identify 437 similar edges across groups, of which 20 are between protein and immune cell phenotypes, as presented in Table 4. Due to the definition of partial correlations based on precision matrices, positive edge strengths correspond to negative conditional dependencies, and vice versa, allowing for the interpretation of overall edge strengths across groups in this context [3].

Table 3 shows the seven differential edges identified by the proposed ONDSA framework, suggesting different compositions of immune cell phenotypes and variations in the conditional dependencies between immune cell pairs across the three *APOE* groups. The immune cells were measured as proportions of live lymphocytes or gated monocytes, following hierarchical structures based on cell subtype definitions. Consequently, some immune cell subtypes have inherent relationships with their parent or sibling cell types. For instance, classical monocytes (CM) and non-classical monocytes (NCM) are sibling subsets of total monocytes (CD14 total), potentially introducing inverse correlations when controlling for CD14 total. Additionally, some immune cell phenotypes are defined as ratios of two cell types, imposing inherent relationships due to this definition. As a result, our observations of a negative dependency between CD8 Cytotoxic T cells and the

CD4/CD8 (CD4 T helper cells / CD8 Cytotoxic T cells) ratio, and between CM and NCM, and a positive dependency between CD8 T naive (Tn) cells and the CD8 Tn/Tm ratio (where CD8 Tm = CD8 Teff + CD8 Tem + CD8 Tcm; Tm, memory T cells; Teff, T effector cells; Tem, T effector-memory cells; Tcm, T central memory cells), and between CD14 total and NCM are consistent with subtype hierarchical definitions but vary across *APOE* groups.

Moreover, immune responses or aging processes can drive the differentiation of Tn cells into more senescent phenotypes [48]. Thus, the different strengths of dependencies across *APOE* groups for these immune cell pairs indicate varying levels of immune activation, potentially contributing to the neuroinflammatory mechanisms underlying cognitive aging. Particularly, our findings of differential conditional dependencies between CD8 Tn and CD8 Tn/Tm ratio show variations in the shift from naive cells to differentiated memory cells across *APOE* groups. In $\epsilon2$ carriers, the positive dependency between CD8 Tn cells and CD8 Tn/Tm ratio has a stronger strength than in the other two groups. This stronger dependency suggests fewer senescent Tm cells in this group, aligning with the potential protective effect of $\epsilon2$ alleles. Similarly, CD4 Tn cells are activated to differentiate into CD4 Tcm cells in adaptive immune responses [49]. As shown in Table 3, differentiation levels vary across *APOE* groups between CD4 Tn and CD4 Tcm cells, with the strongest positive dependency in $\epsilon2$ carriers and the weakest in $\epsilon4$ carriers. Tcm cells are known to have long-term protective effects against pathogens [50], and our results indicate $\epsilon2$ carriers have more persistent CD4 Tcm cells, while $\epsilon4$ carriers have the fewest. Furthermore, we observe differential associations between natural killer (NK) cells and CD38+ T cells, as well as between NK cells and the CD8+Tc17/Treg ratio (Treg refers to CD8 T regulatory cells, CD8+CD25+FoxP3+), with these two interactions being stronger in $\epsilon4$ carriers. NK cells are responsible for the surveillance and elimination of senescent cells [51]. CD38 is an activation marker for T cells, and NK cells possess receptors that recognize CD38 [52], facilitating cell–cell interactions. Activated CD38+ T cells can secrete cytokines that interact with NK cells during immune responses [53, 54], potentially leading to heightened immune activation and inflammation in $\epsilon4$ carriers. Tc17 cells, characterized by the production of pro-inflammatory cytokine IL-17, and CD8 Treg cells, which increase in the blood of older individuals and induce T-cell senescence [55, 56], are both known to interact with NK cells [57], with exacerbated interactions observed in $\epsilon4$ carriers. It is noteworthy that *APOE* is more often discussed in the context of cognitive aging than immunosenescence, and the roles of innate and adaptative immunity in cognitive aging are still largely unknown [58]. Interestingly, the cross-sectional conditional dependency might not fully capture the transition and differentiation of cell types over time. In conclusion, the cell-to-cell interactions identified from differential edges have implications for molecular mechanisms in neuroinflammation. However, interpreting whether the cell-to-cell dependency strengths reflect exacerbated or attenuated neuroinflammation due to *APOE* genotype differences will require further investigation. Additionally, the absence of differential edges between proteins and immune cells is unsurprising, given the limited power for detecting differential edges in imbalanced group designs similar to this sample ($\epsilon2$ carriers account for <15% of the total sample and group sample sizes are relatively modest).

Among the 437 similar edges within or between omics identified by ONDSA (Fig. 5), 250 of them are between protein pairs, 167 of them are between circulating immune cell pairs, and 20 of them are between immune cell and protein pairs. The edges

Table 4. Twenty similar edges across *APOE* genotype groups between protein and circulating immune cell phenotypes identified by the ONDSA framework for the illustrative FHS data

Protein	Protein full name	Circulating immune cell phenotype	Overall edge strength
CD8A	T-cell surface glycoprotein CD8 alpha chain	T cells	−0.48
SCF	Kit ligand	CD4 Tcm	−0.24
TGF-alpha	Protransforming growth factor alpha	CD8 Tn	0.85
TGF-alpha	Protransforming growth factor alpha	CD8 Tn/Tm	−1.74
MCP-4	C-C motif chemokine 13	CD8+ TEMRA	−0.58
MCP-4	C-C motif chemokine 13	CD8+Tc17/Treg	−0.92
FGF-21	Fibroblast growth factor 21	CD38+ T cells	−0.41
FGF-21	Fibroblast growth factor 21	CD38+HLADR+ T cells	0.49
IL-15RA	Interleukin-15 receptor subunit alpha	CD8 Tn/Tm	1.69
IL-10RB	Interleukin-10 receptor subunit beta	CD8+CD25+	−0.38
CXCL5	C-X-C motif chemokine 5	HLA-DR+ T cells	−0.56
TNF	Tumor necrosis factor	CD8+CD27−	−1.39
CD5	T-cell surface glycoprotein CD5	CD4+CD28−CD27−	−0.48
CD5	T-cell surface glycoprotein CD5	CD8+CD28−CD27−	−1.28
CCL3	C-C motif chemokine 3	CD8 Teff	−1.02
INF-gamma	Interferon gamma	CD8+ TEMRA	0.47
MCP-2	C-C motif chemokine 8	CD4 Tn	−1.77
CASP-8	Caspase-8	CD38+ T cells	0.50
CASP-8	Caspase-8	CD38+HLADR+ T cells	−0.68
TNFB	Lymphotoxin-alpha	NK cells	0.30

Note: Positive edge strength indicates negative dependency, while negative edge strength indicates positive dependency. Tcm, T central memory cells; Tn, T naïve cells; Tm, memory T cells; TEMRA cells, T effector memory cells; Tregs cells, CD8 T regulatory cells; Teff, T effector cells; NK cells, natural killer cells.

identified between protein pairs reflect the shared conditional dependency across the full sample, providing a perspective of protein–protein interactions (PPIs). These PPI networks are validated against public databases, such as BioGRID [59] and STRING [60], which include both experimentally confirmed and predicted protein interactions. The majority of identified edges are supported by curated database annotations, co-expression patterns, or co-mentions in scientific literature. Notably, a number of interactions are corroborated by experimental data, such as IL8 and MCP-1 [61], OPG and TRANCE [62], MCP-1 and MCP-4 [61], CCL4 and CCL3 [63], and CD6 and CD5 [64]. The similar edges between circulating immune cell pairs, on the other hand, highlight potential intercellular communication networks across the full sample. As discussed earlier, the immune cells in this study were calculated based on hierarchical structures, reflecting inherent relationships between immune cell subtypes and their parent or sibling types. Many of the interactions identified between immune cell pairs align with well-established patterns of immune communication via cytokines, direct cell-cell contact, and signaling pathways [65, 66]. Thus, the interactions we identified among circulating immune cells are consistent with known immune dynamics and functional hierarchies.

Table 4 focuses on all 20 similar edges identified between immune cell and protein pairs. Some inflammatory biomarkers are known to be secreted by immune cells and can activate immune cells, with these protein biomarkers engaging in intricate interactions with immune cells [67–69]. However, few studies have assessed the associations between the two omics—immune cell phenotypes and inflammatory protein biomarkers—in large community-based samples. To validate the biological relevance of these 20 similar edges identified, we sought both direct and indirect evidence from existing literature. For example, CD8A encodes the alpha chain of the CD8 antigen, a co-receptor on cytotoxic T lymphocytes (CD8+ T cells) that facilitates antigen recognition and T-cell receptor signaling [69]. SCF enhances the function of dendritic cells, which are crucial for T-cell activation [70]. IL-15RA is essential for IL-15 signaling, critical for the
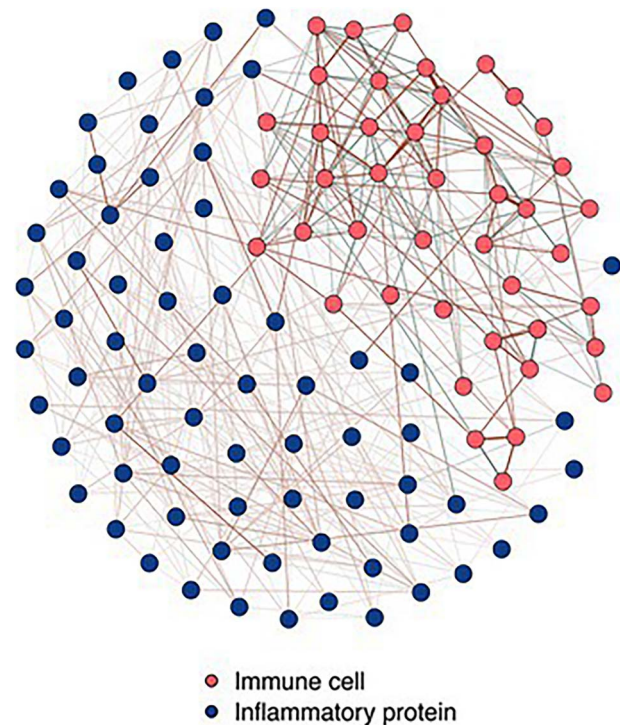

Figure 5. **Similar structures identified by ONSDA in the FHS multi-omics data application.** The red nodes represent the 43 immune cell phenotypes and the black nodes represent the 68 inflammatory protein biomarkers. The strength of edges is represented by the overall conditional dependency across all the three *APOE* genotype groups, where positive dependency is in brown and negative dependency is in green.

homeostasis of memory CD8+ T cells [71]. IL-10RB is part of the receptor complex for IL-10, an anti-inflammatory cytokine broadly expressed by various immune cells [72], including CD8+ T cells that express CD25, a marker of activation [73]. TNF is a pro-inflammatory cytokine that plays a crucial role in immune

regulation, including T-cell activation, proliferation, and apoptosis, through its interaction with TNF receptors (TNFR) [74]. The TNFR family member CD27, expressed by naive human CD8+ T cells, is lost during clonal expansion and differentiation into CD8+CD27− cells as part of this regulatory pathway [75]. CCL3 is produced by activated T cells and recruits additional effector T cells to sites of inflammation [76]. MCP-2 serves as an inflammatory mediator stimulating the directional migration of CD4+ T cells [77]. TNFB can influence the development and function of NK cells [78]. These results support the biological plausibility of the conditional associations identified by ONDSA across the full sample.

Our findings contribute to the understanding of neuro-inflammation and cognitive aging by revealing how *APOE* genotypes influence immune cell interactions and their associations with inflammatory proteins. The differential edges identified among immune cell pairs across *APOE* groups suggest that the *APOE* genotype affects the communication and conditional dependencies between circulating immune cells (Table 3). These variations may reflect differences in immune activation or regulation that contribute to neuroinflammatory processes associated with cognitive aging. The similar edges—including those between immune cells and proteins—that are consistent across *APOE* groups may indicate common pathways that influence inflammatory protein levels (Table 4). Our previous work demonstrated that certain inflammatory proteins, such as TGF-alpha, MCP-4, and FGF-21, are associated with executive function or visuospatial cognitive domain scores, and higher levels of TGF-alpha, TNF, CCL3, and TNFB are linked to increased risks of incident dementia or AD, with some associations varying by *APOE* genotype [47]. By integrating these findings, we propose that *APOE* genotypes modulate immune cell interactions, which in turn influence inflammatory protein levels, ultimately affecting cognitive aging outcomes. The identification of these network structures provides insight into the mechanisms by which immune dysregulation may contribute to neuroinflammation and cognitive decline in an *APOE* genotype-dependent manner. However, due to the imbalanced group design and limited sample size, replication studies with larger samples are needed to further validate these potential biological pathways.

Discussion

In this paper, we propose a novel statistical testing framework, ONDSA, for identifying differential and similar structures targeting multiple omics networks, represented by sparse GGMs, in a two-step procedure that effectively controls FDRs at a desired level. By addressing the limitations of existing methods that are restricted to two-group comparisons or lack theoretical guarantees, ONDSA enables rigorous statistical tests across precision matrices from multiple conditions. This advancement enhances our ability to uncover meaningful inter- and intra-omics pathways associated with disease onset and progression. Ren *et al.* [16] developed a tuning-free estimator for the precision matrix of a single GGM and proved its asymptotic normality and efficiency; however, non-sparse estimators were obtained for sparse GGMs. This creates a tradeoff between asymptotic properties and interpretability since dense networks make it difficult to evaluate GGMs across multiple groups without formal comparisons. ONDSA builds on the estimators from Ren *et al.* [16] to sequentially estimate differential and similar omics network

structures while adjusting for the correlated effects of entry-wise comparisons in the FDR control procedures. Other existing methods for joint estimation of multiple sparse GGMs lack rigorous tests to compare across groups. Our proposed ONDSA demonstrates superior performance, achieving a better balance between TPR and FPR than existing methods.

Simulation studies demonstrate that ONDSA's performance varies for different graph structures in balanced groups. When the common graph structures are band or hub graphs, sparseness assumptions might fail, making structures of interest harder to detect. Limited sample sizes in high-dimensional settings can lead to slight FDR inflation and reduced power However, when the common graph is scale-free, FDR is well controlled, and power remains at reasonably high levels across balanced scenarios. Recent studies on omics networks such as metabolomic, transcriptomic, and proteomic networks have shown that their topologies belong to scale-free networks [79]. Our simulation studies for balanced groups show that ONDSA performs particularly well when the common graph follows a scale-free structure. Thus, by applying ONDSA to high-throughput omics data, we provide a valuable tool for integrating large-scale multi-omics data and identifying pathways regulating biological mechanisms across different subgroups.

We further demonstrate through simulation studies that ONDSA significantly reduces computational costs compared with alternative methods. ONDSA leverages a computationally efficient and tuning-free GGM estimation approach, feasible for handling tens of thousands of variables [16, 35, 39]. We acknowledge that as the number of variables in omics networks increases, the number of statistical tests will also increase, which could impact the computational time of ONDSA. However, this challenge can be effectively mitigated using parallel computing techniques.

The limitations of ONDSA arise in imbalanced group settings, particularly with extremely modest sample size groups, where caution is required. Simulations show inflated FDR and limited power for identifying differential omics network structures in low-dimensional and high-dimensional settings, though inflated FDR is less pronounced in the latter. In imbalanced designs, some differential edges may be misclassified as similar due to minor group differences or limited sample size, and conclusions on similar structures are more reliable in low-dimensional settings. Our real data application on FHS multi-omics data highlights the challenges posed by small and imbalanced group sizes, such as in the *APOE* $\epsilon 2$ carriers group.

To summarize, several practical limitations should be considered when applying ONDSA. As mentioned earlier, ONDSA may be sensitive to highly imbalanced group sizes (with groups comprising <15% of the total sample), particularly when these groups have extremely modest sample sizes. In such cases, caution is required when interpreting results, especially for differential edges in low-dimensional settings and both differential and similar structures in high-dimensional settings. The minimum sample size per group should scale proportionally with the number of variables p. Following Ren *et al.* [16], we recommend that the group sample size n grows as p increases, particularly in high-dimensional settings, to ensure consistent estimation of precision matrices and reliable use of ONDSA. Additionally, ONDSA relies on a sparsity assumption in the precision matrix estimator [16]. When applied to highly connected networks, such as hub or band graphs, the accuracy in identifying differential edges may decrease, though the method remains effective for identifying similar structures. Users should be

mindful of this limitation when applying ONDSA to denser networks. Furthermore, it is important to note that ONDSA is based on GGMs, where conditional independence is equivalent to a zero precision matrix entry for multivariate normal data [3]. If this normality assumption is violated, such equivalence breaks down, and conditional independence may not correspond to a zero entry in the precision matrix. This complexity is beyond the scope of this work, and researchers are encouraged to consult relevant literature on conditional dependency for non-normal data (e.g. [80]) to ensure the appropriate use of ONDSA, especially when normal transformations are not feasible. Future research could investigate extending our proposed methodologies to mixed graphical models, which accommodate both Gaussian and multinomial variables.

This ONDSA framework also shows great promise in identifying shared mechanisms across ancestry groups, analyzing brain connectivity based on functional magnetic resonance imaging data across individuals, and investigating interactions among single-cell omics data across tissues. Recently, a debiasing procedure with entry-wise confidence intervals based on asymptotic normality is proposed [34], which can be coupled with many other existing precision matrix estimators. Since one of the essential prerequisites for constructing the proposed test statistics is the use of a GGM estimator with asymptotic normality properties, the ONDSA framework offers a truly unified approach. This framework is not only designed for the specific estimator we employ [16], but is also generalizable to several other state-of-the-art de-sparsified or debiased precision matrix estimators [17, 18, 34]. This flexibility makes ONDSA a powerful and versatile tool for high-dimensional omics network analysis, broadening its applicability across a range of cutting-edge methodologies.

Key Points

- We propose a rigorous statistical testing framework, ONDSA, for differential and similarity analyses targeting multiple omics networks represented by sparse GGMs.
- ONDSA addresses a critical gap in multi-group precision matrix comparisons with proper FDR control, allowing us to identify meaningful inter- and intra-omics pathways associated with disease onset and progression.
- The robustness to various realistic conditions, combined with its computational efficiency, makes ONDSA a superior option for researchers aiming to analyze complex omics networks efficiently and accurately.
- We provide a valuable tool for integrating large-scale multi-omics data and have developed an efficient and well-documented R package for ONDSA, available at: https://github.com/jiachenchen322/ONDSA.

Acknowledgments

Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Funding

Data and code availability

Framingham Heart Study Offspring data are available for request via NHLBI Biological Specimen and Data Repositories Information Coordinating Center (BioLINCC). The proposed ONDSA framework is developed into an R package available at: https://github.com/jiachenchen322/ONDSA.

Competing interests

The authors declare no competing interests.

Author contributions

J.C. performed conceptualization, methodology development and implementation, numerical simulations, real data curation and analysis, wrote the original draft, and did revisions and editing. K.L.L. and J.M.M. performed conceptualization, funding acquisition, real data curation, and investigation, wrote the original draft, and did revisions and editing.

References

1. Altenbuchinger M, Weihs A, Quackenbush J. *et al.* Gaussian and mixed graphical models as (multi-) omics data analysis tools. *Biochim Biophys Acta, Gene Regul Mech* 2020;**1863**:194418. https://doi.org/10.1016/j.bbagrm.2019.194418.
2. Lauritzen SL. *Graphical Models*. Oxford: Oxford University Press, 1996. https://doi.org/10.1093/oso/9780198522195.001.0001.
3. Shutta KH, De Vito R, Scholtens DM. *et al.* Gaussian graphical models with applications to omics analyses. *Stat Med* 2022;**41**: 5150–87. https://doi.org/10.1002/sim.9546.
4. Meinshausen N, Bühlmann P. High-dimensional graphs and variable selection with the lasso. *Ann Stat* 2006;**34**:1436–62. https://doi.org/10.1214/009053606000000281.
5. Yuan M, Lin Y. Model selection and estimation in the Gaussian graphical model. *Biometrika* 2007;**94**:19–35. https://doi.org/10.1093/biomet/asm018.
6. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 2008;**9**:432–41. https://doi.org/10.1093/biostatistics/kxm045.
7. d'Aspremont A, Banerjee O, El Ghaoui L. First-order methods for sparse covariance selection. *SIAM J Matrix Anal Appl* 2008;**30**: 56–66. https://doi.org/10.1137/060670985.
8. Rothman AJ, Bickel PJ, Levina E. *et al.* Sparse permutation invariant covariance estimation. *Electron J Stat* 2008;**2**:494–515. https://doi.org/10.1214/08-EJS176.
9. Peng J, Wang P, Zhou N. *et al.* Partial correlation estimation by joint sparse regression models. *J Am Stat Assoc* 2009;**104**:735–46. https://doi.org/10.1198/jasa.2009.0126.
10. Fan J, Feng Y, Wu Y. Network exploration via the adaptive LASSO and SCAD penalties. *Ann Appl Stat* 2009;**3**:521–41. https://doi.org/10.1214/08-AOAS215.

11. Yuan M. High dimensional inverse covariance matrix estimation via linear programming. *J Mach Learn Res* 2010;**11**: 2261–86.

12. Cai T, Liu W, Luo X. A constrained L1 minimization approach to sparse precision matrix estimation. *J Am Stat Assoc* 2011;**106**: 594–607. https://doi.org/10.1198/jasa.2011.tm10155.

13. Ravikumar P, Wainwright MJ, Raskutti G. *et al.* High-dimensional covariance estimation by minimizing L1-penalized log-determinant divergence. *Electron J Stat* 2011;**5**:935–80. https://doi.org/10.1214/11-EJS631.

14. Liu H, Han F, Yuan M. *et al.* High-dimensional semiparametric Gaussian copula graphical models. *Ann Stat* 2012;**40**:2293–326. https://doi.org/10.1214/12-AOS1037.

15. Liu W. Gaussian graphical model estimation with false discovery rate control. *Ann Stat* 2013;**41**:2948–78. https://doi.org/10.1214/13-AOS1169.

16. Ren Z, Sun T, Zhang CH. *et al.* Asymptotic normality and optimalities in estimation of large Gaussian graphical models. *Ann Stat* 2015;**43**:991–1026. https://doi.org/10.1214/14-AOS1286.

17. Janková J, van de Geer S. Confidence intervals for high-dimensional inverse covariance estimation. *Electron J Stat* 2015;**9**: 1205–29. https://doi.org/10.1214/15-EJS1031.

18. Janková J, van de Geer S. Honest confidence regions and optimality in high-dimensional precision matrix estimation. *Test* 2017;**26**:143–62. https://doi.org/10.1007/s11749-016-0503-5.

19. Ni Y, Stingo FC, Baladandayuthapani V. Sparse multi-dimensional graphical models: A unified Bayesian framework. *J Am Stat Assoc* 2017;**112**:779–93. https://doi.org/10.1080/01621459.2016.1167694.

20. Ma J, Karnovsky A, Afshinnia F. *et al.* Differential network enrichment analysis reveals novel lipid pathways in chronic kidney disease. *Bioinformatics* 2019;**35**:3441–52. https://doi.org/10.1093/bioinformatics/btz114.

21. Danaher P, Wang P, Witten DM. The joint graphical lasso for inverse covariance estimation across multiple classes. *J R Stat Soc Series B Stat Methodology* 2014;**76**:373–97. https://doi.org/10.1111/rssb.12033.

22. Shojaie A. Differential network analysis: A statistical perspective. *Wiley Interdisciplinary Reviews: Computational Statistics* 2021;**13**:e1508. https://doi.org/10.1002/wics.1508.

23. Guo J, Levina E, Michailidis G. *et al.* Joint estimation of multiple graphical models. *Biometrika* 2011;**98**:1–15. https://doi.org/10.1093/biomet/asq060.

24. Zhang B, Wang Y. Learning structural changes of Gaussian graphical models in controlled experiments. In: Grünwald P, and Spirtes P. (eds). *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, UAI 2010, pp. 701–708. AUAI Press, 2010. https://event.cwi.nl/uai2010/papers.html.

25. Ha MJ, Baladandayuthapani V, Do KA. DINGO: Differential network analysis in genomics. *Bioinformatics* 2015;**31**:3413–20. https://doi.org/10.1093/bioinformatics/btv406.

26. Cai T, Li H, Liu W. *et al.* Joint estimation of multiple high-dimensional precision matrices. *Statistica Sinica* 2016;**26**:445–64. https://doi.org/10.5705/ss.2014.256.

27. Zhao SD, Cai T, Li H. Direct estimation of differential networks. *Biometrika* 2014;**101**:253–68. https://doi.org/10.1093/biomet/asu009.

28. Haslbeck JM. Estimating group differences in network models using moderation analysis. *Behav Res Methods* 2022;**54**:522–40. https://doi.org/10.3758/s13428-021-01637-y.

29. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodology* 1995;**57**:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

30. Xia Y, Cai T, Cai T. Testing differential networks with applications to the detection of gene-gene interactions. *Biometrika* 2015;**102**:247–66. https://doi.org/10.1093/biomet/asu074.

31. He H, Cao S, Zhang J. *et al.* A statistical test for differential network analysis based on inference of Gaussian graphical model. *Sci Rep* 2019;**9**:10863. https://doi.org/10.1038/s41598-019-47362-7.

32. Liu W. Structural similarity and difference testing on multiple sparse Gaussian graphical models. *Ann Stat* 2017;**45**:2680–707. https://doi.org/10.1214/17-AOS1539.

33. Li KC, Palotie A, Yuan S. *et al.* Finding disease candidate genes by liquid association. *Genome Biol* 2007;**8**:1–9.

34. Li S, Cai T, Li H. Transfer learning in large-scale gaussian graphical models with false discovery rate control. *J Am Stat Assoc* 2023;**118**:2171–83. https://doi.org/10.1080/01621459.2022.2044333.

35. Zhang R, Ren Z, Chen W. SILGGM: An extensive R package for efficient statistical inference in large-scale gene networks. *PLoS Comput Biol* 2018;**14**:e1006369. https://doi.org/10.1371/journal.pcbi.1006369.

36. Anderson TW. *An Introduction to Multivariate Statistical Analysis*, 3rd Edition. New York: Wiley, 2003.

37. Sun T, Zhang CH. Comment: Minimax estimation of large covariance matrices under L1 norm. *Statist Sinica* 2012;**22**:1354–8.

38. Efron B. Correlation and large-scale simultaneous significance testing. *J Am Stat Assoc* 2007;**102**:93–103. https://doi.org/10.1198/016214506000001211.

39. Wang T, Ren Z, Ding Y. *et al.* FastGGM: An efficient algorithm for the inference of Gaussian graphical model in biological networks. *PLoS Comput Biol* 2016;**12**:e1004755. https://doi.org/10.1371/journal.pcbi.1004755.

40. Zhang R, Ren Z, Chen W. SILGGM: An extensive R package for efficient statistical inference in large-scale gene networks. *PLoS Comput Biol* 2018;**14**:e1006369. https://doi.org/10.1371/journal.pcbi.1006369.

41. Zhao T, Liu H, Roeder K. *et al.* The huge package for high-dimensional undirected graph estimation in R. *J Mach Learn Res* 2012;**13**:1059–62.

42. Feinleib M, Kannel WB, Garrison RJ. *et al.* The Framingham offspring study. *Design and preliminary data Preventive Medicine* 1975;**4**:518–25. https://doi.org/10.1016/0091-7435(75)90037-7.

43. Fang Y, Doyle MF, Chen J. *et al.* Circulating immune cell phenotypes are associated with age, sex, CMV, and smoking status in the Framingham heart study offspring participants. *Aging (Albany NY)* 2023;**15**:3939–66. https://doi.org/10.18632/aging.204686.

44. Raulin AC, Doss SV, Trottier ZA. *et al.* ApoE in Alzheimer's disease: Pathophysiology and therapeutic strategies. *Mol Neurodegener* 2022;**17**:72. https://doi.org/10.1186/s13024-022-00574-4.

45. Soares HD, Potter WZ, Pickering E. *et al.* Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease. *Arch Neurol* 2012;**69**:1310–7. https://doi.org/10.1001/archneurol.2012.1070.

46. Ho GJ, Drego R, Hakimian E. *et al.* Mechanisms of cell signaling and inflammation in Alzheimer's disease. *Curr Drug Targets* 2005;**4**:247–56. https://doi.org/10.2174/1568010053586237.

47. Chen J, Doyle MF, Fang Y. *et al.* Peripheral inflammatory biomarkers are associated with cognitive function and dementia: Framingham heart study offspring cohort. *Aging Cell* 2023;**22**:e13955. https://doi.org/10.1111/acel.13955.

48. Mittelbrunn M, Kroemer G. Hallmarks of T cell aging. *Nat Immunol* 2021;**22**:687–98. https://doi.org/10.1038/s41590-021-00927-z.

49. Gasper DJ, Tejera MM, Suresh M. CD4 T-cell memory generation and maintenance. *Crit Rev Immunol* 2014;**34**:121–46. https://doi.org/10.1615/CritRevImmunol.2014010373.

50. Gehad A, Teague JE, Matos TR. *et al.* A primary role for human central memory cells in tissue immunosurveillance. *Blood Adv* 2018;**2**:292–8. https://doi.org/10.1182/bloodadvances.2017011346.

51. Kim K, Admasu TD, Stolzing A. *et al.* Enhanced co-culture and enrichment of human natural killer cells for the selective clearance of senescent cells. *Aging (Albany NY)*. 2022;**14**:2131–47. https://doi.org/10.18632/aging.203931.

52. Piedra-Quintero ZL, Wilson Z, Nava P. *et al.* CD38: An immunomodulatory molecule in inflammation and autoimmunity. *Front Immunol* 2020;**11**:597959. https://doi.org/10.3389/fimmu.2020.597959.

53. Li W, Liang L, Liao Q. *et al.* CD38: An important regulator of T cell function. *Biomed Pharmacother* 2022;**153**:113395. https://doi.org/10.1016/j.biopha.2022.113395.

54. Ghosh A, Khanam A, Ray K. *et al.* CD38: An ecto-enzyme with functional diversity in T cells. *Front Immunol* 2023;**14**:1146791. https://doi.org/10.3389/fimmu.2023.1146791.

55. Jagger A, Shimojima Y, Goronzy JJ. *et al.* Regulatory T cells and the immune aging process: A mini-review. *Gerontology* 2014;**60**:130–7. https://doi.org/10.1159/000355303.

56. Ye J, Huang X, Hsueh EC. *et al.* Human regulatory T cells induce T-lymphocyte senescence. *Blood* 2012;**120**:2021–31. https://doi.org/10.1182/blood-2012-03-416040.

57. Kale A, Sharma A, Stolzing A. *et al.* Role of immune cells in the removal of deleterious senescent cells. *Immun Ageing* 2020;**17**:16. https://doi.org/10.1186/s12979-020-00187-9.

58. Fernández-Calle R, Konings SC, Frontiñán-Rubio J. *et al.* APOE in the Bullseye of neurodegenerative diseases: Impact of the APOE genotype in Alzheimer's disease pathology and brain diseases. *Mol Neurodegener* 2022;**17**:62. https://doi.org/10.1186/s13024-022-00566-4.

59. Stark C, Breitkreutz BJ, Reguly T. *et al.* BioGRID: A general repository for interaction datasets. *Nucleic Acids Res* 2006;**34**:D535–9. https://doi.org/10.1093/nar/gkj109.

60. Szklarczyk D, Kirsch R, Koutrouli M. *et al.* The STRING database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 2023;**51**:D638–46. https://doi.org/10.1093/nar/gkac1000.

61. Von Hundelshausen P, Agten SM, Eckardt V. *et al.* Chemokine interactome mapping enables tailored intervention in acute and chronic inflammation. *Sci Transl Med* 2017;**9**:eaah6650. https://doi.org/10.1126/scitranslmed.aah6650.

62. Yasuda H, Shima N, Nakagawa N. *et al.* Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL. *Proc Natl Acad Sci* 1998;**95**:3597–602. https://doi.org/10.1073/pnas.95.7.3597.

63. Guan E, Wang J, Norcross MA. Identification of human macrophage inflammatory proteins 1*α* and 1*β* as a native secreted heterodimer. *J Biol Chem* 2001;**276**:12404–9. https://doi.org/10.1074/jbc.M006327200.

64. Gimferrer I, Farnós M, Calvo M. *et al.* The accessory molecules CD5 and CD6 associate on the membrane of lymphoid T cells. *J Biol Chem* 2003;**278**:8564–71. https://doi.org/10.1074/jbc.M209591200.

65. Frankenstein Z, Alon U, Cohen IR. The immune-body cytokine network defines a social architecture of cell interactions. *Biol Direct* 2006;**1**:1–15. https://doi.org/10.1186/1745-6150-1-32.

66. Olbei M, Thomas JP, Hautefort I. *et al.* CytokineLink: A cytokine communication map to analyse immune responses–case studies in inflammatory bowel disease and COVID-19. *Cells* 2021;**10**:2242. https://doi.org/10.3390/cells10092242.

67. Chen L, Deng H, Cui H. *et al.* Inflammatory responses and inflammation-associated diseases in organs. *Oncotarget* 2018;**9**:7204–18. https://doi.org/10.18632/oncotarget.23208.

68. Furman D, Campisi J, Verdin E. *et al.* Chronic inflammation in the etiology of disease across the life span. *Nat Med* 2019;**25**:1822–32. https://doi.org/10.1038/s41591-019-0675-0.

69. Murphy K, Weaver C. *Janeway's Immunobiology.* New York: Garland Science, 2016. 10.1201/9781315533247.

70. Broudy VC. Stem cell factor and hematopoiesis. *Blood* 1997;**90**:1345–64.

71. Waldmann TA. The biology of interleukin-2 and interleukin-15: Implications for cancer therapy and vaccine design. *Nat Rev Immunol* 2006;**6**:595–601. https://doi.org/10.1038/nri1901.

72. Saraiva M, O'garra A. The regulation of IL-10 production by immune cells. *Nat Rev Immunol* 2010;**10**:170–81. https://doi.org/10.1038/nri2711.

73. Herndler-Brandstetter D, Schwaiger S, Veel E. *et al.* CD25-expressing CD8+ T cells are potent memory cells in old age. *J Immunol* 2005;**175**:1566–74. https://doi.org/10.4049/jimmunol.175.3.1566.

74. Croft M. The role of TNF superfamily members in T-cell function and diseases. *Nat Rev Immunol* 2009;**9**:271–85. https://doi.org/10.1038/nri2526.

75. Ochsenbein AF, Riddell SR, Brown M. *et al.* CD27 expression promotes long-term survival of functional effector–memory CD8+ cytotoxic T lymphocytes in HIV-infected patients. *J Exp Med* 2004;**200**:1407–17. https://doi.org/10.1084/jem.20040717.

76. Schaller TH, Batich KA, Suryadevara CM. *et al.* Chemokines as adjuvants for immunotherapy: Implications for immune activation with CCL3. *Expert Rev Clin Immunol* 2017;**13**:1049–60. https://doi.org/10.1080/1744666X.2017.1384313.

77. Taub DD, Proost P, Murphy WJ. *et al.* Monocyte chemotactic protein-1 (MCP-1),-2, and-3 are chemotactic for human T lymphocytes. *J Clin Invest* 1995;**95**:1370–6. https://doi.org/10.1172/JCI117788.

78. Allan DS, Rybalov B, Awong G. *et al.* TGF-*β* affects development and differentiation of human natural killer cell subsets. *Eur J Immunol* 2010;**40**:2289–95. https://doi.org/10.1002/eji.200939910.

79. Chen H, Sharp BM. Content-rich biological network constructed by mining PubMed abstracts. *BMC Bioinformatics* 2004;**5**:1–13. https://doi.org/10.1186/1471-2105-5-147.

80. Baba K, Shibata R, Sibuya M. Partial correlation and conditional correlation as measures of conditional independence. *Aust N Z J Stat* 2004;**46**:657–64. https://doi.org/10.1111/j.1467-842X.2004.00360.x.