

RESEARCH

Open Access



Non parametric differential network analysis: a tool for unveiling specific molecular signatures

Pietro Hiram Guzzi¹, Arkaprava Roy², Marianna Milano^{3*} and Pierangelo Veltri⁴

*Correspondence:
m.milano@unicz.it

¹ Department of Medical and Surgical Sciences, Magna Graecia University, Catanzaro, Italy

² University of Florida, Gainesville, FL, USA

³ Department of Experimental and Clinical Medicine, Magna Graecia University, Catanzaro, Italy

⁴ Department of Computer Science, Modelling and Electronics DIMES, University of Calabria, Rende, Italy

Abstract

Background: The rewiring of molecular interactions in various conditions leads to distinct phenotypic outcomes. Differential network analysis (DINA) is dedicated to exploring these rewirings within gene and protein networks. Leveraging statistical learning and graph theory, DINA algorithms scrutinize alterations in interaction patterns derived from experimental data.

Results: Introducing a novel approach to differential network analysis, we incorporate differential gene expression based on sex and gender attributes. We hypothesize that gene expression can be accurately represented through non-Gaussian processes. Our methodology involves quantifying changes in non-parametric correlations among gene pairs and expression levels of individual genes.

Conclusions: Applying our method to public expression datasets concerning diabetes mellitus and atherosclerosis in liver tissue, we identify gender-specific differential networks. Results underscore the biological relevance of our approach in uncovering meaningful molecular distinctions.

Keywords: Differential network, DINA, Differential network analysis, Atherosclerosis, Diabetes

Introduction

The emergence of high-throughput technologies in genomics, proteomics, and non-coding RNA studies has revolutionized our understanding of how variations in the abundance of these biological molecules correlate with diseases [1]. This deluge of data has spurred the creation of innovative analytical techniques that adopt a system-level approach through the lens of network science. These techniques employ networks to depict the complex interactions between biological molecules under specific conditions, as inferred from empirical data [2]. A particularly compelling use of network theory is comparing biological networks under disparate conditions, such as contrasting a disease state with a healthy one.

Differential network analysis (DINA) has gained recognition for its ability to delineate the differences between two states by encapsulating them within a singular differential network that highlights their variances [3–6]. DINA has found application in contrasting various experimental conditions or phenotypes, with recent studies underscoring the



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

significance of factors like the age and sex of patients on drug response, disease progression, and comorbidity prevalence in chronic diseases [7–9]. Empirical findings have demonstrated notable disparities in disease incidence and progression based on sex and age, such as the heightened susceptibility of older diabetic patients to comorbidities and the observed sex-based differences in COVID-19 mortality rates [10–12]. This underscores the need for advanced algorithms to unravel the molecular mechanisms driving these age- and sex-dependent disparities.

DINA algorithms are designed to pinpoint changes in network structures by identifying association measures that differ between two biological states, C_1, C_2 . When presented with two disparate biological conditions, represented by two networks of molecular interactions, DINA algorithms aim to uncover the network rewiring that underpins the mechanistic differences between these states. DINA algorithms construct networks $\mathcal{N}_1, \mathcal{N}_2$ for each condition, starting with gene expression datasets from two conditions. These networks feature nodes for each gene and weighted edges that denote the strength and nature of associations or causal relationships between genes. A differential network \mathcal{N}_d is then derived to represent the variation in associations across conditions, a technique previously applied to investigate disease-related alterations [8, 13].

Traditional methods often assume that gene expression data adhere to specific parametric distributions, such as Gaussian or Poisson distributions [1, 14–16]. However, the count-based nature of Next-Generation Sequencing (NGS) data challenges these assumptions, prompting the necessity for non-parametric DINA analysis approaches.

We introduce a novel DINA methodology that identifies differential edges between networks and integrates differential gene expressions, taking into account sex-based differences [12]. This approach employs multivariate count data for predicting gene expression levels and constructs a conditional dependence graph using pairwise Markov random fields [17]. This departure from traditional methods, which often presuppose parametric distributions for gene expression data, highlights the imperative for non-parametric techniques in DINA analysis.

Our proposed DINA algorithm is initiated by constructing two condition-specific graphs, from which a final differential graph is derived. This graph is then pruned to emphasize edges related to genes exhibiting differential expression. Our DINA method facilitates the identification of differential networks while incorporating considerations of gender differences, thereby advancing our understanding of the molecular basis of disease in the context of sex and age.

Relevance of the DNA

Differential networks are used to compare two or more networks based on changes in connectivity or interactions between nodes under different conditions or across different datasets. They are commonly applied in biological networks, such as gene expression or protein-protein interaction networks, to identify how interactions change in response to disease, treatment, or other stimuli [18–26].

For instance, DNA may be applied on Expression Networks in Healthy vs. Diseased States. Let imagine that Network 1 represents Healthy State. In the healthy state, a set of genes (nodes) are connected based on their co-expression (edges) For example, if Gene A and Gene B are highly co-expressed, there would be an edge between them. This

network shows how genes typically interact in a healthy individual. Let imagine a Network 2 who models a diseased state (such as cancer) in which the expression levels of genes may change, and therefore the interactions between them may be altered. Some gene-gene interactions may be lost, and new interactions may emerge. The differential network is created by taking the difference between the two networks (Healthy vs. Diseased). This network will highlight both *lost and novel connections*. Interactions present in the healthy network but missing in the diseased network, and interactions that appear in the diseased state but were absent in the healthy state.

For example: In the healthy state: Gene A interacts with Gene B and Gene C. In the diseased state: Gene A no longer interacts with Gene B, but interacts with Gene D instead. The differential network would show a lost edge between Gene A and Gene B and a new edge between Gene A and Gene D. This type of analysis helps in identifying key genes and pathways that may be responsible for the diseased condition or that could be potential therapeutic targets see Fig. 1.

Examples of the relevance of DNA are reported in many works. For instance, by Ha et al [27], who describes differential networks between two different subtypes of glioblastoma estimated from genomic data. Basha et al., [28] introduces an extensive differential network analysis of multiple human tissue interactomes. As a results they are able to evidence differences of processes between tissues.

Related work

DINA's application in distinguishing differentially expressed genes among various sample groups is invaluable, especially in contrasting individuals with specific diseases against healthy controls. This methodological approach is crucial in molecular biology and bioinformatics for pinpointing genes with variable expression levels between diseased and healthy sample groups.

Central to DINA-based research are algorithms designed to detect alterations in network structures under varying conditions [29]. These algorithms have been pivotal in

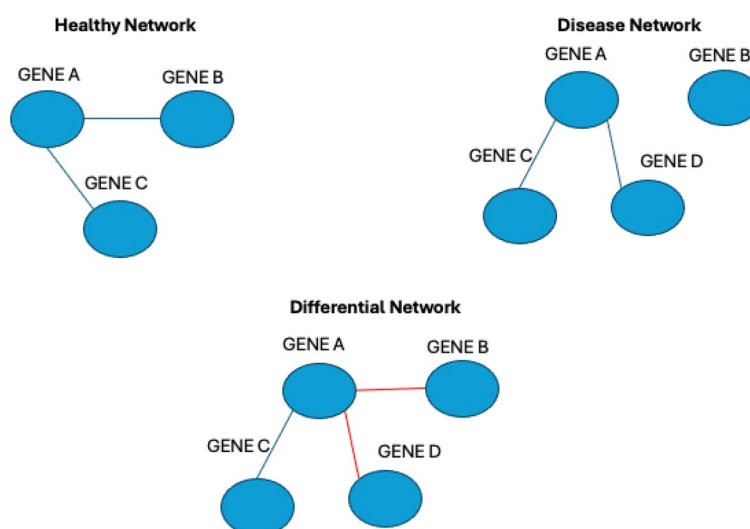


Fig. 1 Toy example of differential network

biology for mapping the transition from healthy to diseased states within the same biological framework [2]. Our focus narrows to networks that maintain constant node sets yet exhibit variable edge sets. Specifically, in the presence of two distinct conditions \mathcal{C}_1 and \mathcal{C}_2 , represented by graphs $G_1(V, E_1)$ and $G_1(V, E_2)$, the objective of DINA analysis is to pinpoint the modifications of the network.

In biological systems analysis, it is pertinent to note that while nodes represent directly quantifiable entities, the derivation of edges necessitates observing a sequence of temporal data. For instance, gene networks originating from microarray experiments necessitate the inference of edges from data through *statistical graphical models* [30–32]. In these models, each node within the graph $G = (V, E)$ is one of these measurable random variables X_1, \dots, X_M , and the edges quantify a pre-specified notion of associations between the pairs of these variables. In this setting, the focus is predominantly on undirected graphs where the directions or the causality of these associations are not of interest. Among different metrics of associations, partial correlation is one of the most common ones as it measures conditional dependencies. Probabilistic graphical models allow conditional dependency-based graph estimation. Differential associations within these models are scrutinized by evaluating the variance in partial correlations across experimental conditions, utilizing specific statistical tests to measure the alterations in correlations among entities. Additionally, changes in gene expression levels are assessed using the classical Student's t-test [33]. Subsequently, these statistical evaluations are amalgamated into a singular optimization model aiming to elucidate the hierarchical network structures. However, certain assumptions inherent to previous models, such as Gaussian data distribution, may not hold across all experimental conditions, necessitating non-parametric methods. While computationally efficient and simpler to implement, these methods demand adherence to specific distributional prerequisites, failing which could skew or invalidate the results.

Several studies have opted for a nonparanormal data distribution (or Gaussian copula) approach [34], employing rank-based correlation matrices like Spearman correlation or Kendall's τ . There are other variations available too [29]. However, the nonparanormal models are primarily only suitable for continuous data, which limits its applicability in other settings.

These models have found applications in analyzing brain data and sequencing counts, circumventing the temporal limitations of non-parametric methods. Efficient Bayesian models have emerged [29], calculating edge probabilities by inferring their likelihood. Some methods adopt diverse heuristics for probability inference, challenging direct data derivation, as highlighted in [17], with this method surpassing other contemporary techniques.

Non-parametric methods, recognized for their minimal assumptions regarding data distribution, leverage data-driven approaches to evaluate network connectivity differences between conditions. Their flexibility and robustness are advantageous in handling complex, non-linear node relationships within networks, albeit at the cost of computational intensity and reduced interpretability.

The decision between parametric and non-parametric approaches for differential network analysis hinges on data characteristics, foundational assumptions, and the investigative query. Researchers frequently use sensitivity analysis and result cross-validation

to ensure their findings' robustness and reliability. Integrating insights from both methodological spectrums can yield a more detailed comprehension of the differential network architecture.

Materials and methods

Non parametric differential network analysis algorithm

Let us consider two different expression datasets encoded in two matrices $N_j \times M$ (N_j samples, M genes) for $j = 1, 2$ denoted as X^1, X^2 representing two biological conditions C_1, C_2 . Each row of X^j stores the expression values of M genes of different samples. Therefore, $X_{i,j}^c$ ($c = 1, 2, i = 1, \dots, N_c, j = 1, \dots, M$) denotes expression of j -th gene in i -th sample under condition c . Note that the sample sizes under the two conditions may be different. We model this data under a Bayesian non-parametric framework.

Each column representing a gene may be encoded as a network node and compute conditional independence-based graphical relation. Let $M \times M$ dimensional matrices \mathcal{P}_1 and \mathcal{P}_2 represents the conditional independence relation among the M genes [17]. We then define the differential relation between two conditions based on the posterior samples of \mathcal{P}_1 and \mathcal{P}_2 .

Following the pairwise Markov random field (MRF) model for counts from [17], we consider the following joint probability mass function for M -dimensional count-valued data X ,

$$\Pr(X_1, \dots, X_M) \propto \exp \left(\sum_{j=1}^M [\alpha_j X_j - \log(X_j!)] - \sum_{\ell=2}^M \sum_{j<\ell} \beta_{j\ell} F(X_j) F(X_\ell) \right),$$

where $F(\cdot)$ is a monotone increasing bounded function with support $[0, \infty)$. As in [17], we let $F(\cdot) = (\tan^{-1}(\cdot))^\theta$ for some positive $\theta \in \mathbb{R}^+$. Since the data is positive-valued, the range of $F(X)$ is $(0, (\frac{\pi}{2})^\theta)$ and the exponent θ is specified as a minimizer of the loss, quantifying the difference in covariance between $F(X)$ and X following [17]. For detailed descriptions of the method, readers are encouraged to check [17].

Under this model, if $\beta_{j\ell} = 0$, we have X_j and X_ℓ to be conditionally independent, i.e. $P(X_j, X_\ell | X_{-(j,\ell)}) = P(X_j | X_{-(j,\ell)})P(X_\ell | X_{-(j,\ell)})$, where $X_{-(j,\ell)}$ stands for all the variables excluding X_j and X_ℓ . Our estimated graphical relation thus would rely on $\beta_{j\ell}$'s.

We take a Bayesian route for inference and put the same priors as in [17]. Specifically, for the $\beta_{j\ell}$'s, we set simple independent and identically distributed mean zero Gaussian priors. The parameter λ_j 's are treated as random effects and given distributions D_j . The distribution D_j governs the over-dispersion and the shape of the marginal count distribution for the j^{th} node. To allow these marginals to be flexibly determined by the data, we take a Bayesian nonparametric approach using Dirichlet process (DP) priors, where $D_j \sim \text{DP}(M_j D_0)$, with D_0 as a Gamma base measure and M_j as a precision parameter. The precision parameter M_j follows a Gamma distribution, $M_j \sim \text{Ga}(c, d)$, allowing for greater adaptivity to the data. We run MCMC to approximate the posterior and generate posterior samples of $\beta_{j\ell}$'s. The model follows a structure, similar to the Poisson auto-model [35]. When $\beta_{j\ell}$'s are zero, the marginals lead to Poisson-type marginals with sample-specific means as λ_j 's are modeled as random effects. Thus, they also need to be sampled for all the samples. This puts a moderately high computational cost. The general

code to fit this model is in the second author's GitHub page <https://github.com/royarkaprava/CONGA>.

After running the Markov chain Monte Carlo (MCMC) sampling for the above model under two conditions, we get the matrices \mathcal{P}_1 and \mathcal{P}_2 with (j, k) -th entries as $\beta_{j,k}^{(1)}$ and $\beta_{j,k}^{(2)}$ respectively. Consequently, a differential network is defined as the difference $\beta_{j,k}^{(1)} - \beta_{j,k}^{(2)}$ for each edge (j, k) where $\beta_{j,k}^{(1)}$ and $\beta_{j,k}^{(2)}$ are the coefficients under two conditions 1 and 2. From the MCMC samples, we can get the posterior mean of these differences as $\hat{\beta}_{j,k}^{(1)} - \hat{\beta}_{j,k}^{(2)}$ using the individual posterior means. Alternatively, we can compute other posterior summaries such as $P(|\beta_{j,k}^{(1)} - \beta_{j,k}^{(2)}| > c \mid D_1, D_2)$, which is the posterior probability that $|\beta_{j,k}^{(1)} - \beta_{j,k}^{(2)}|$ is greater than some pre-specified cutoff c given the two datasets, denoted as D_1 and D_2 . We take the second approach to define our differential networks. To choose c adequately, we run a sensitivity test. Specifically, we vary c over a range, and compute $f_{ij}(c) = P(|\beta_{j,k}^{(1)} - \beta_{j,k}^{(2)}| > c \mid D_1, D_2)$ for each choice. Then, we monitor $\sum_{j,k} (f_{ij}(c_1) - f_{ij}(c_2))^2$ and find the smallest c showing stability around its neighborhood to assess its sensitivity on the estimate.

Databases

The T2DiACoD database, as described by Rani et al. (2017) [36], was employed to collate a comprehensive list of genes linked to comorbid conditions associated with Type 2 Diabetes Mellitus (T2DM). Gene expression datasets were also sourced from the GTEx database [37].

T2DiACoD is a meticulously compiled database, the result of rigorous research and systematic literature review. It catalogues genes and noncoding RNAs that are crucial to understanding T2DM and its frequent comorbidities, including atherosclerosis, nephropathy, diabetic retinopathy, and cardiovascular disorders. This repository, enriched through meticulous data integration from existing databases, encapsulates 650 genes and 34 microRNAs related to these conditions, providing a reliable resource for your research.

The genotype-tissue expression (GTEx) project is a vast open-access platform that empowers researchers like us with the distribution of genomic data collected from various individuals. This repository encompasses a broad spectrum of genomic data, from sequencing to methylation analyses, providing a wealth of information at your fingertips.

GTEx offers detailed metadata for each sample, covering aspects such as tissue type, sex, and age, categorized into six distinct groups. This makes GTEx an invaluable asset for investigating the interplay between age and tissue-specific gene expression. As of February 1st, the GTEx database boasts a collection of 17,382 samples across 54 tissue types from 948 donors, all accessible via the GTEx web portal. This portal enables users to efficiently search for and visualize data [12, 38, 39]. Furthermore, the data can be downloaded for in-depth analysis with custom scripts. In our research, we leveraged data from various tissues, dividing samples into two categories based on sex. Our study focused on the genes that play a pivotal role in developing T2DM-related complications, examining nine tissues (including blood, brain, adipose, amygdala, aorta, colon, coronary, liver, and lung tissues).

Table 1 Characteristics of the differential networks related to diabetes

DN	Nodes	Edges
Liver	128	3340
Aorta	237	2116
Heart	238	4316

Table 2 Characteristics of the differential networks related to atherosclerosis

DN	Nodes	Edges
Adipose Visceral	12	32
Artery Coronary	11	30
Artery Tibial	11	12
Blood	2	1

To ensure the validity of our research, we meticulously selected an equal number of samples from each tissue type, maintaining a balanced representation of age groups and uniformity in sample sizes across tissues. This approach facilitated an equitable distribution of age groups within each sex-based category for each tissue analyzed, ensuring the precision and reliability of our findings.

Focusing on atherosclerosis, we obtained a list of 115 genes related to this disease in T2DiACoD database, while for diabetes we obtained a list of 650 genes. We retrieved expression data by employing GTExVisualizer [8, 40], and metadata related to tissue, sex and age of the sample are extracted using genes identified in the T2DiACoD database in the previous step. Expression data are measured as Transcript per Million (TPM). This data integration and gene enrichment process was performed using an ad-hoc realised script that has been integrated into GTExVisualizer. We performed the analysis at tissue level, thus for each considered tissue, we split the data into male and female samples and randomly selected the same number of samples. We first generated DN by using non parametric methods and we evaluated the biological significance by means of enrichment methods.

Results

To show the effectiveness of our method we present two case studies on two chronic diseases to show differential mechanisms related to sex differences. We evaluated differential networks between men and women focusing on genes related to diabetes and atherosclerosis as reported in the T2DiaCoD database [36]. The characteristics of all the DNs are reported in Tables 1 and 2. Then, to evaluate the biological significance of the resulted DNs, we performed a pathway enrichment analysis. Functional enrichment analysis, also known as pathway enrichment analysis (PEA), is a bioinformatics technique used to identify biological pathways that are significantly over-represented in a given list of genes compared to what would be expected by chance. These biological functions are stored in bioinformatics databases such KEGG, and statistical methods like Fisher's exact test, that computes the p-value of the enrichment are used to determine the most enriched pathways. All PEA tools require a gene list as input.

Thus, for each network, we performed a pathway based on the KEGG pathway database [41] available on STRING enrichment app [42] of the Cytoscape software [43].

Liver Tissue For liver tissue in diabetes, we obtained a DN with 128 nodes and 3340 edges, see Fig. 2. The enrichment analysis highlighted the presence of some enriched pathways between sex, see Table 3. Figure 3 depicts the subnetworks of the differential network for Liver tissue in Diabetes, where, we highlighted the genes involved in the resulted pathways in different colours. In order to evaluate

Heart Tissue For Hearth tissue in Diabetes we obtained a DN with 238 nodes and 4316 edges, see Fig. 6. The enrichment analysis highlighted the presence of some enriched pathways between sex, see Table 5. Figure 7 depicts the subnetworks of the differential network for Liver tissue in Diabetes, where, we highlighted the genes involved in the resulted pathways in different colours.

[illegible]

Fig. 2 The figure shows the differential network between men and women focusing on genes in liver tissue related to diabetes disease

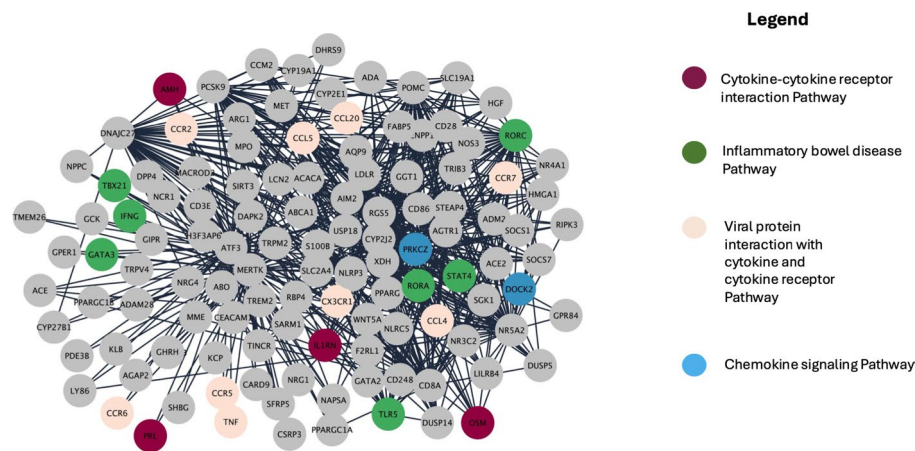


Fig. 3 The figure shows selected subnetworks of the differential network for Liver tissue in Diabetes. In detail, starting from the differential network, we highlighted the genes involved in the pathways resulted from enriched analysis in different colours, whereas the genes not involved in pathways are identified in grey. In particular, the genes involved in Cytokine-cytokine receptor interaction are coloured in purple, the genes involved in Inflammatory bowel disease are coloured in green, the genes involved in Viral protein interaction with cytokine and cytokine receptor are coloured in pink and the genes involved in Chemokine signaling pathway are coloured in blue

Table 3 Top Enriched Pathways of liver tissue in Diabetes

Description	Term name	P-value with Fisher
Cytokine-cytokine receptor interaction	hsa04060	5.15e ⁻⁸
Viral protein interaction with cytokine and cytokine receptor	hsa04061	3.48e ⁻⁷
Chemokine signaling pathway	hsa04062	3.47e ⁻⁶
Inflammatory bowel disease	hsa05321	3.47e ⁻⁶
Toll-like receptor signaling pathway	hsa04620	5.97e ⁻⁵
Malaria	hsa05144	2.6e ⁻⁴
Rheumatoid arthritis	hsa05323	2.6e ⁻⁴
Renin-angiotensin system	hsa04614	3.5e ⁻⁴
Th17 cell differentiation	hsa04659	4.1e ⁻⁴
Chagas disease	hsa05142	4.1e ⁻⁴

Finally, we applied the Markov Clustering Algorithm to perform a cluster analysis on differential networks. Figures 9, 10, 11 depict the clustered differential networks for Liver, Aorta, Heart tissues.

Atherosclerosis related differential networks

Adipose Visceral Tissue For Adipose Visceral tissue in Atherosclerosis we obtained a DN with 12 nodes and 32 edges, see Fig. 12. The enrichment analysis did not report the presence of some enriched pathways between sex.

Artery Coronary Tissue For Artery Coronary tissue in Atherosclerosis we identified a network of 11 nodes and 30 edges, see Fig. 13. The enrichment analysis highlighted the presence of some enriched pathways between sex, see Table 6. Figure 14 depicts

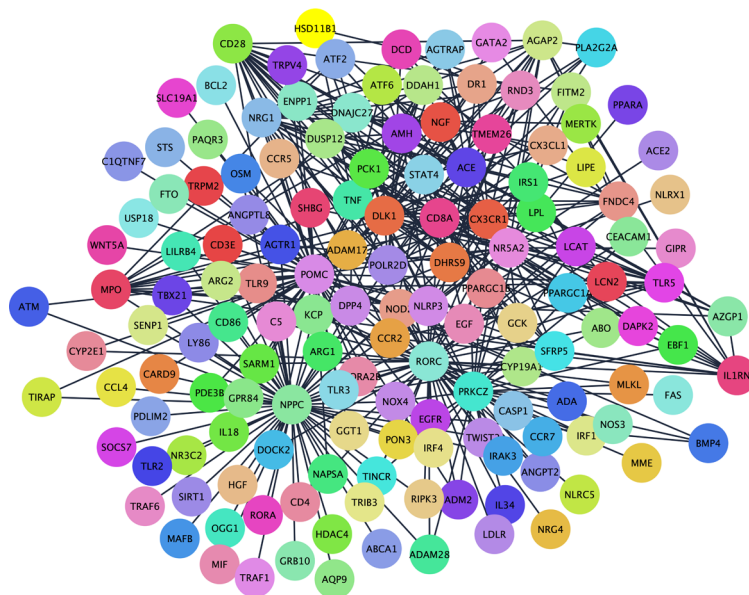


Fig. 4 The figure shows the differential network between men and women focusing on genes in aorta tissue related to diabetes disease

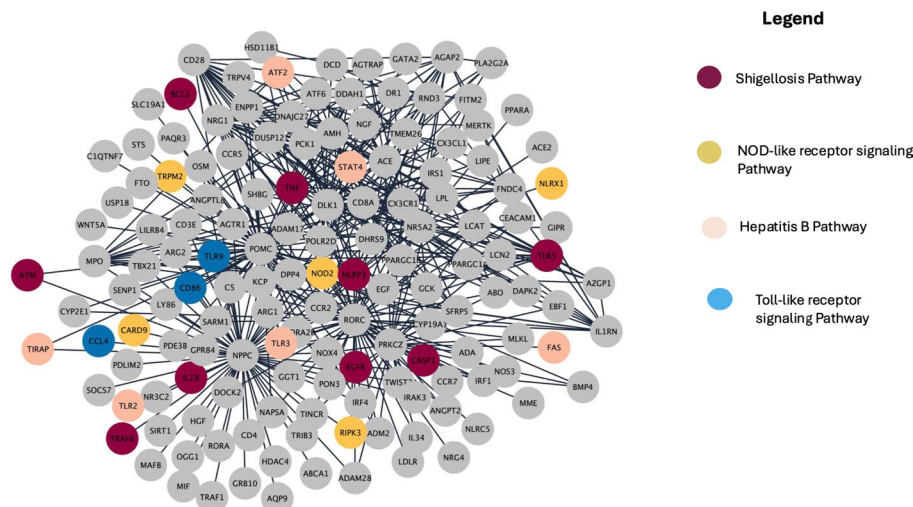


Fig. 5 The figure shows Selected subnetwork of the differential network for Aorta tissue in Diabetes. In detail, starting from the differential network, we highlighted the genes involved in the pathways resulted from enriched analysis in different colours, whereas the genes not involved in pathways are identified in grey. In particular, the genes involved Toll-like receptor signaling pathway are coloured in blue, the genes involved in NOD-like receptor signaling pathway are coloured in yellow, the genes involved in Hepatitis B are coloured in pink, and the genes involved in Shigellosis are coloured in purple

the subnetworks of the differential network for Liver tissue in Diabetes, where, we highlighted the genes involved in the resulted pathways in different colours.

Artery Tibial Tissue For Artery Tibial tissue in Atherosclerosis we identified a network of 11 nodes and 12 edges, see Fig. 15. The enrichment analysis highlighted the presence of some enriched pathways between sex, see Table 4. Figure 16 depicts the

Table 4 Top Enriched Pathways of aorta tissue in Diabetes

Description	Term name	P-value with Fisher
Toll-like receptor signaling pathway	hsa04620	$7.05e^{-19}$
NOD-like receptor signaling pathway	hsa04621	$7.57e^{-16}$
Shigellosis	hsa05131	$7.57e^{-16}$
Hepatitis B	hsa05161	$1.08e^{-15}$
Salmonella infection	hsa05132	$1.5e^{-15}$
TNF signaling pathway	hsa04668	$4.27e^{-15}$
NF-kappa B signaling pathway	hsa04064	$1.31e^{-14}$
Toxoplasmosis	hsa05145	$1.55e^{-14}$
PD-L1 expression and PD-1 checkpoint pathway in cancer	hsa05235	$1.82e^{-14}$
PI3K-Akt signaling pathway	hsa04151	$6.36e^{-14}$

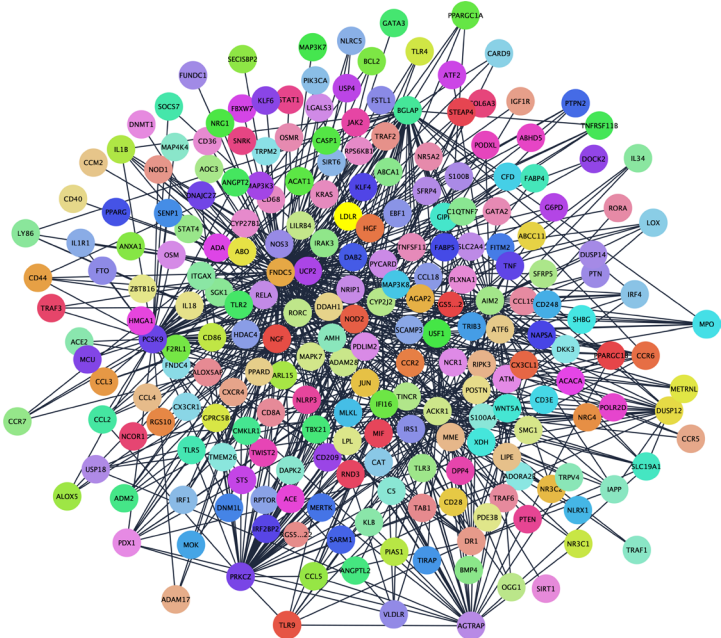


Fig. 6 The figure shows the differential network between men and women focusing on genes in heart tissue related to diabetes disease

subnetwork of the differential network for Liver tissue in Diabetes, where, we highlighted the genes involved in the resulted pathways in different colours.

Blood Tissue For Blood tissue in Atherosclerosis we identified a network of 2 nodes and 1 edges, see Fig. 17. The enrichment analysis did not report the presence of some enriched pathways between sex.

We summarize in Fig. 18 a selected subnetwork of the differential network in Adipose Visceral, Artery Coronary, and Aorta Tissue, Artery Tibial and Blood tissues.

Finally, we applied the Markov Clustering Algorithm to perform a cluster analysis on differential networks. Figures 19, 20, 21 depict the clustered differential networks for Adipose Visceral, Artery Coronary, Artery Tibial tissues.

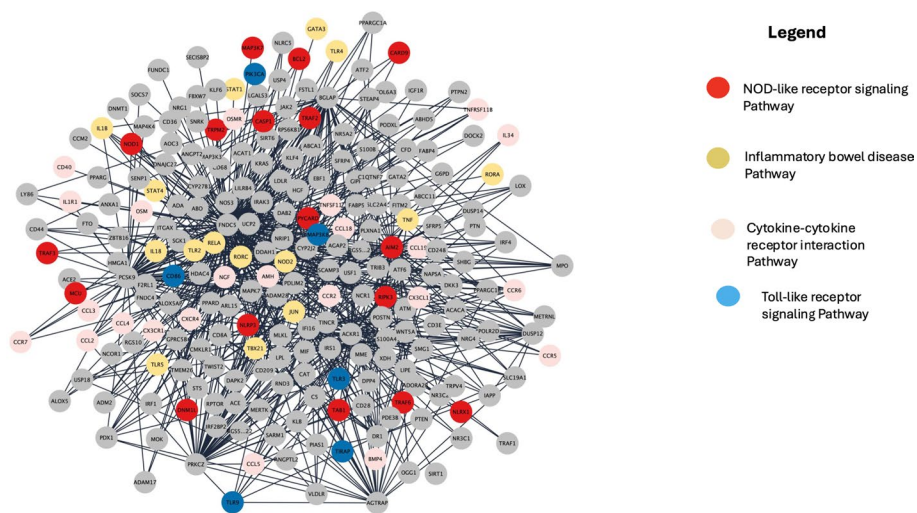


Fig. 7 The figure shows selected subnetwork of the differential network for Heart tissue in Diabetes. In detail, starting from the differential network, we highlighted the genes involved in the pathways resulted from enriched analysis in different colours, whereas the genes not involved in pathways are identified in grey. In particular, the genes involved Toll-like receptor signaling pathway are coloured in blue, the genes involved in NOD-like receptor signaling pathway are coloured in red, the genes involved in Cytokine-cytokine receptor interaction are coloured in pink, and the genes involved in Inflammatory bowel disease are coloured in yellow

Table 5 Top Enriched Pathways of heart tissue in Diabetes

Description	Term name	P-value with Fisher
NOD-like receptor signaling pathway	hsa04621	2.33e ⁻¹⁸
Toll-like receptor signaling pathway	hsa04620	8.48e ⁻¹⁸
TNF signaling pathway	hsa04668	6.15e ⁻¹⁶
Cytokine-cytokine receptor interaction	hsa04060	4.69e ⁻¹⁴
NF-kappa B signaling pathway	hsa04064	2.34e ⁻¹³
Salmonella infection	hsa05132	3.34e ⁻¹³
Inflammatory bowel disease	hsa05321	3.79e ⁻¹³
Shigellosis	hsa05131	5.73e ⁻¹³
PD-L1 expression and PD-1 checkpoint pathway in cancer	hsa05235	3.01e ⁻¹²
Viral protein interaction with cytokine and cytokine receptor	hsa04061	1.04e ⁻¹¹

Comparison with baseline methods

In order to evaluate the effectiveness of our method with respect to other parametric ones, we compared Conga with a R package iDINGO [44]. iDINGO is able to infer group-specific dependencies and build differential networks. We built the differential network for diabetes dataset and atherosclerosis dataset. We run iDINGO using default parameters. However, iDINGO built the differential networks according to NP-complete approach by including in the building of the network the whole set of edges. We reported the characteristics of the differential networks related to diabetes and atherosclerosis in Tables 7 and 8.

We subsequently performed the path enrichment analysis on the DN, and found that the pathways detected for each tissue for diabetes were fewer than those built

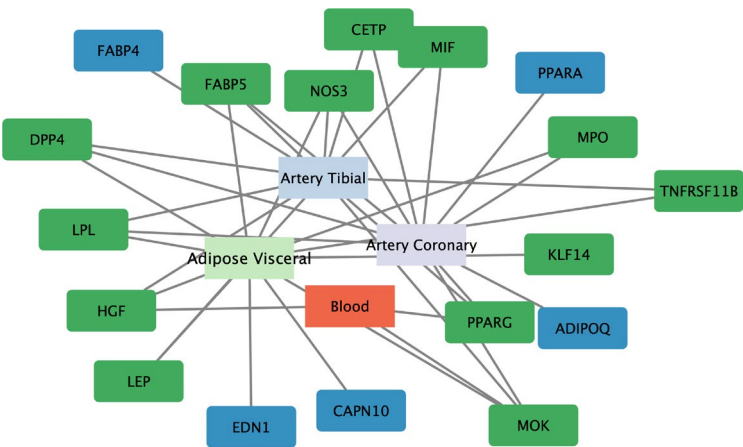


Fig. 8 Differential network related to diabetes expressed in Liver, Heart, Aorta tissues. The genes expressed in different tissues are reported in green, whereas the genes of specific tissue are reported in blue

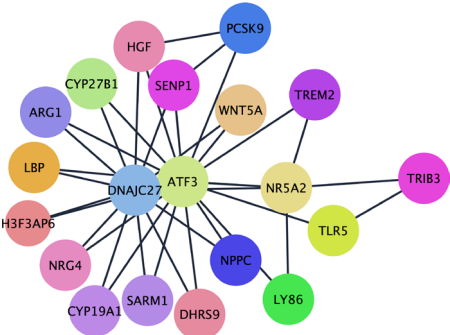


Fig. 9 Clustered Differential subnetwork for Liver tissue in Diabetes

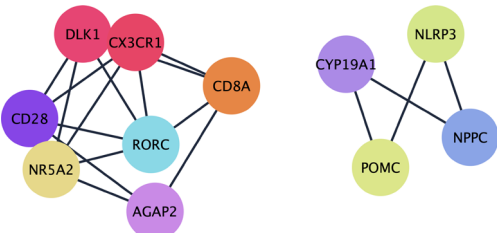


Fig. 10 Clustered Differential subnetwork for Aorta tissue in Diabetes

with our method, whereas none pathway was detected for atherosclerosis. This demonstrates that our methodology is able to build networks that have a greater biological information content than classical methods. We reported the results of pathway enrichment analysis in Tables 9, 10 and 11.

Discussion

In our research study, we introduced a method called differential network analysis. This approach is particularly beneficial for datasets that contain count data or non-parametric data, which are prevalent in biological research but pose challenges due to

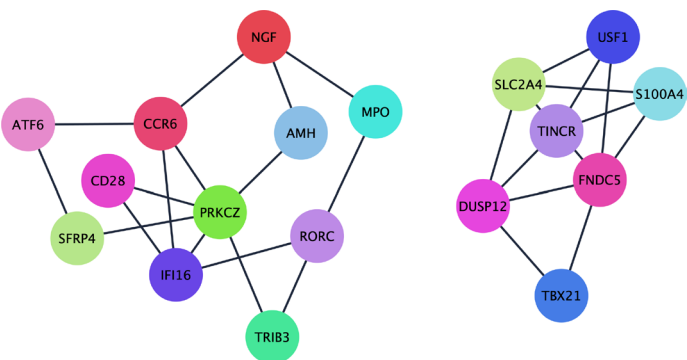


Fig. 11 Clustered Differential subnetwork for Heart tissue in Diabetes

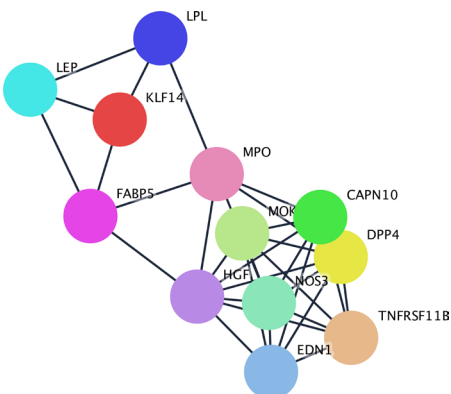


Fig. 12 Differential network for Adipose Visceral tissue in Atherosclerosis

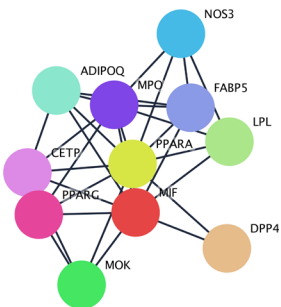


Fig. 13 The figure shows the differential network between men and women focusing on genes in artery coronary tissue related to atherosclerosis disease

their non-normal distribution and discrete nature. By employing differential network analysis, we can delve into the underlying network structures between different conditions, providing a more comprehensive understanding of biological variations and interactions that are not discernible through traditional methods. Given the absence of an established standard in this area, the rigorous evaluation of the biological relevance of the networks derived from our differential network analysis was paramount. We analysed the structure of the networks, ensuring their biological significance and consistency with known biological pathways and mechanisms. This comprehensive

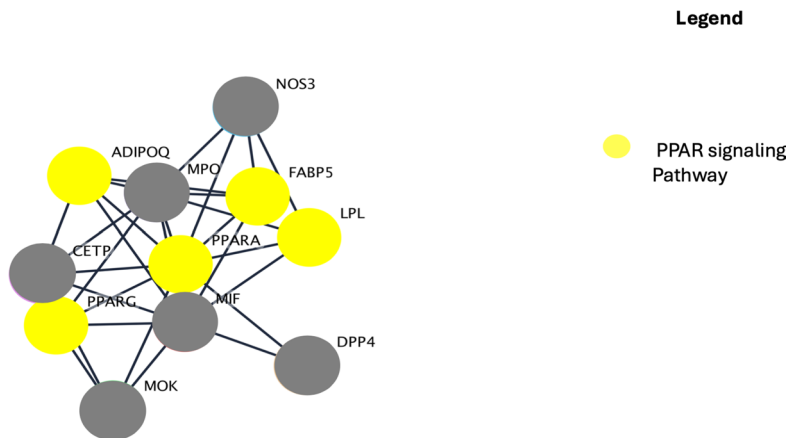


Fig. 14 The figure shows selected subnetwork of the differential network for Artery Coronary tissue in Atherosclerosis. In detail, starting from the differential network, we highlighted the genes involved in the pathways resulted from enriched analysis in yellow, whereas the genes not involved in pathways are identified in grey. In particular, the genes involved in PPAR signaling pathway are reported in yellow

Table 6 Enriched Pathway of Artery Coronary and Artery Tibial tissue in Atherosclerosis

Description	Term name	P-value with Fisher
PPAR signaling pathway of Artery Coronary	hsa03320	1.48e ⁻⁷
PPAR signaling pathway of Artery Tibial	hsa03320	2.59e ⁻⁵

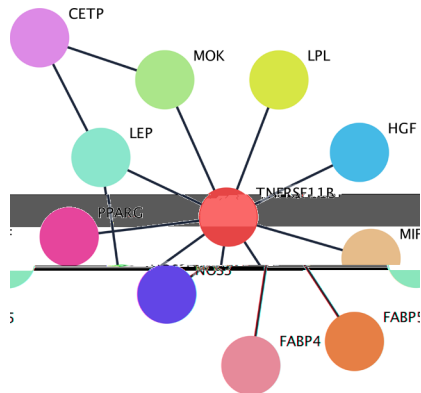


Fig. 15 The figure shows the differential network between men and women focusing on genes in artery tibial tissue related to atherosclerosis disease

analysis not only validated the practical utility of our methodology but also provided novel insights into the biological systems under study. By correlating our findings with existing biological knowledge and experimental data, we improved the biological relevance of our results, further underscoring the value of differential network analysis in uncovering biological patterns and interactions. In the context of diabetes, we found some differential networks in Liver, Aorta, and Heart tissues are enriched for the Toll-like receptor signaling pathway, as indicated in Tables 3, 4, 7. The toll-like receptor 4 (TLR-4) pathway has been associated with various pathophysiological conditions,

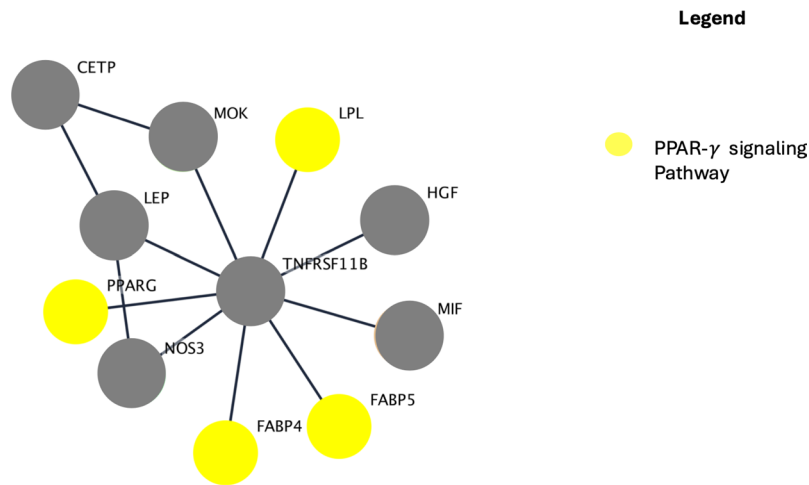


Fig. 16 A Selected subnetwork of the differential network for Artery Tibial tissue in Atherosclerosis. In detail, starting from the differential network, we highlighted the genes involved in the pathways resulted from enriched analysis in yellow, whereas the genes not involved in pathways are identified in grey. In particular, the genes involved in PPAR- γ signaling pathway are reported in yellow

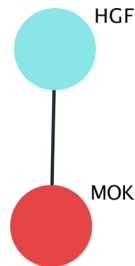


Fig. 17 Differential network for Blood tissue in Atherosclerosis

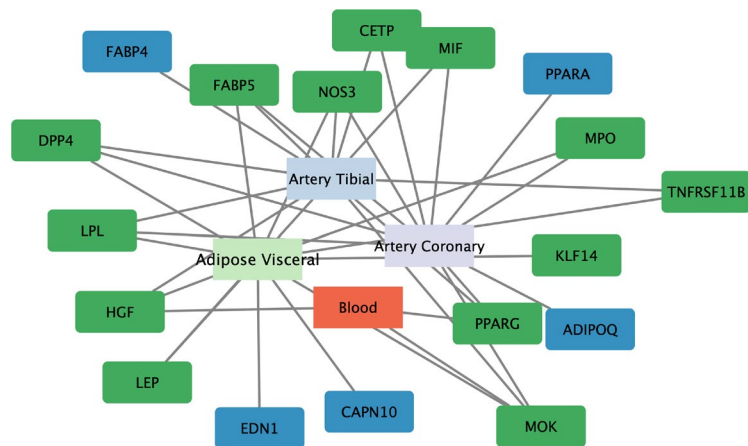


Fig. 18 Differential network related to atherosclerosis expressed in Adipose Visceral, Artery Coronary, Artery Tibial tissues. The genes expressed in different tissues are reported in green, whereas the genes of specific tissue are reported in blue

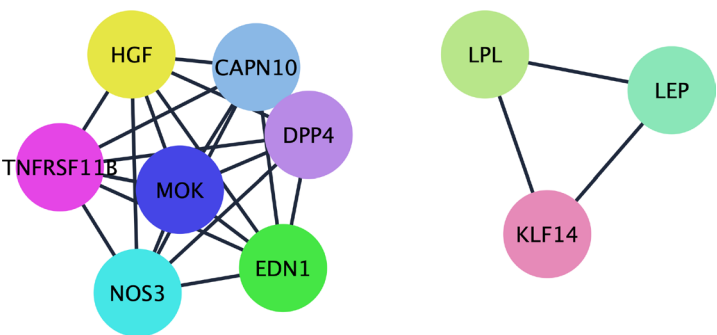


Fig. 19 Clustered Differential network for Adipose Visceral tissue in Atherosclerosis

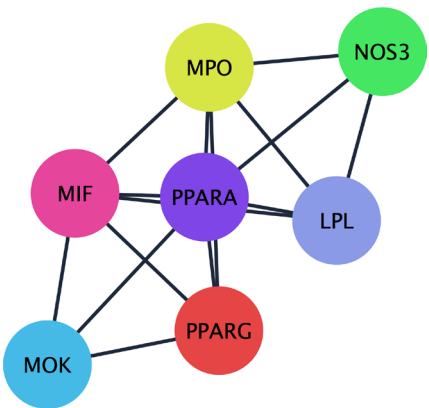


Fig. 20 Clustered Differential network for Artery Coronary tissue in Atherosclerosis

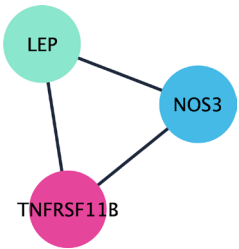


Fig. 21 Clustered Differential network for Artery Tibial tissue in Atherosclerosis

Table 7 Characteristics of the differential networks related to diabetes built with iDINGO

DN	Nodes	Edges
Liver	128	8127
Aorta	238	28679
Heart	238	28679

including cardiovascular diseases (CVDs) and Rheumatoid Arthritis (RA), underscor- ing the relevance of our findings in the broader context of human health. Different studies demonstrated that TLR4 activates the expression of several pro-inflammatory

Table 8 Characteristics of the differential networks related to atherosclerosis built with iDINGO

DN	Nodes	Edges
Adipose Visceral	12	45
Artery Coronary	11	45
Artery Tibial	11	45
Blood	2	1

Table 9 Top Enriched Pathways of liver tissue in Diabetes related to DNs built with iDINGO

Description	Term name	P-value with Fisher
Cytokine-cytokine receptor interaction	hsa04060	$2.31e^{-6}$
Viral protein interaction with cytokine and cytokine receptor	hsa04061	$2.7e^{-6}$
Chemokine signaling pathway	hsa04062	$4.57e^{-5}$
Inflammatory bowel disease	hsa05321	$2.31e^{-6}$
Toll-like receptor signaling pathway	hsa04620	$3.46e^{-4}$
Malaria	hsa05144	$7e^{-4}$

Table 10 Top Enriched Pathways of aorta tissue in Diabetes related to DNs built with iDINGO

Description	Term name	P-value with Fisher
Salmonella infection	hsa05132	$6.12e^{-14}$
TNF signaling pathway	hsa04668	$8.55e^{-14}$
NF-kappa B signaling pathway	hsa04064	$2.24e^{-13}$
Toxoplasmosis	hsa05145	$2.65e^{-13}$
PD-L1 expression and PD-1 checkpoint pathway in cancer	hsa05235	$2.68e^{-13}$
PI3K-Akt signaling pathway	hsa04151	$3.5e^{-12}$

Table 11 Top Enriched Pathways of heart tissue in Diabetes related to DNs built with iDINGO

Description	Term name	P-value with Fisher
NOD-like receptor signaling pathway	hsa04621	$2.25e^{-18}$
Toll-like receptor signaling pathway	hsa04620	$1.17e^{-17}$
TNF signaling pathway	hsa04668	$7.96e^{-16}$
Cytokine-cytokine receptor interaction	hsa04060	$6.43e^{-14}$
NF-kappa B signaling pathway	hsa04064	$2.91e^{-13}$

cytokine genes that play pivotal roles in myocardial inflammation, particularly myocarditis, myocardial infarction, ischemia-reperfusion injury, and heart failure [45–48].

Also, we found in Aorta and Heart tissues the NOD-like receptor signalling pathway; see Tables 6 and 7. The NOD-like Receptor (NLR) family of proteins is a group of pattern recognition receptors (PRRs) known to mediate the initial innate immune response to cellular injury and stress. Different studies reveal the role of the activation

of the Nod-like receptor protein 3 (NLRP3) inflammasome in the pathogenesis of many metabolic diseases, including diabetes and its complications [49].

Furthermore, the differential networks in the Aorta and Heart tissues were enriched for the TNF signalling pathway; see Tables 6 and 7. Lamki et al. [50] reported that Tumor necrosis factor (TNF) represent a central mediator of a broad range of biological activities from cell proliferation, cell death and differentiation to induction of inflammation and immune modulation. TNF mediates the inflammatory response and regulates immune function. Inappropriate production of TNF or sustained activation of TNF signalling has been implicated in the pathogenesis of a broad spectrum of human diseases, including diabetes, cancer, osteoporosis, allograft rejection, and autoimmune diseases such as multiple sclerosis, rheumatoid arthritis, and inflammatory bowel diseases [50, 51].

Instead, for the differential networks in Liver tissue were reported the chemokine signalling pathway that promotes changes in cellular morphology [52] and insulin signalling pathway that accounts for selective insulin resistance [53], see Table 6.

Also, the differential network in the Aorta tissue was enriched for the PI3K-Akt signalling pathway, see Table 7. Phosphatidylinositol 3-kinases (PI3Ks) are crucial coordinators of intracellular signalling in response to extracellular stimulators. The hyperactivation of PI3K signalling cascades is one of the most common events in human cancers. The high recurrence of phosphoinositide 3-kinase (PI3K) pathway adjustments in cancer has led to a surge in the progression of PI3K inhibitors [54]. For diabetes, we found that differential networks in Artery Coronary and Artery Tibial tissue were enriched for the PPAR signalling pathway, which activation is linked to a correlation between metabolic syndromes and cancer, see Table 6.

Conclusion

Differential network analysis (DINA) may help the understanding the intricate interactions within biological systems, especially regarding specific conditions or phenotypes. This study utilizes DINA to explore the differential molecular interactions related to diabetes mellitus, taking into account age and gender-specific variations. Differential networks are critical since they enable the detection and visualisation of variations in gene and protein interactions between males and females. This approach helps to identify subtle biological discrepancies that conventional analytical methods could miss. Non-parametric methods for DINA are emphasized in the study since they enhance the robustness and applicability of the findings, given the complex nature of biological data, rather than assuming normal distributions of gene expression data.

The study applies this method to liver tissue gene expression data related to diabetes, identifying distinct networks that may be involved in sex-specific disease mechanisms. These findings can provide insight into why certain diseases exhibit different manifestations in men and women, ultimately leading to more targeted approaches to treatment and management. Identifying gender-specific differential networks helps to understand the molecular basis of diabetes and its variation with sex, providing potential pathways for therapeutic intervention and a deeper understanding of the etiology of the disease.

This study demonstrates the effectiveness of the differential network approach in discriminating between male and female biological samples. The method effectively

identifies and visualizes differences in molecular interactions related to diabetes in liver tissue between the two sexes. Such discrimination is vital for personalized medicine, leading to more precise diagnostic and therapeutic strategies. By effectively highlighting the differences in gene expression and interactions, the study supports the potential of DINA in improving our understanding of sex-specific traits in diseases, which is crucial for advancing gender-specific medicine.

The article discusses the use of differential networks in biological research, especially in complex diseases such as diabetes. Integrating non-parametric methods improves the analysis by accounting for the inherent complexities and variations in biological data often overlooked in parametric approaches. This research contributes significantly to our understanding of diabetes and sets a precedent for future studies exploring other complex diseases with potential variations in expression and interaction patterns across different groups.

Abbreviations

DINA	Differential network analysis
NGS	Next-generation sequencing
GTE _x	Genotype-tissue expression

Acknowledgements

This work was funded by the Next Generation EU - Italian NRRP, Mission 4, Component 2, Investment 1.5, call for the creation and strengthening of 'Innovation Ecosystems', building 'Territorial R&D Leaders' (Directorial Decree n. 2021/3277) - project Tech4You - Technologies for climate change adaptation and quality of life improvement, n. ECS0000009. This work reflects only the authors' views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them. We acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 9 - Green-aware AI, under the NRRP MUR program funded by the NextGenerationEU.

Author contributions

Conceptualization, P.H.G., R.A. M.M.; Data curation, P.H.G., R.A. M.M and P.V.; Investigation, P.H.G., R.A. M.M and P.V.; Methodology, P.H.G., R.A. M.M and P.V.; writing—original draft preparation, P.H.G., R.A. M.M and P.V.; writing—review and editing, P.H.G., R.A. M.M and P.V. All authors reviewed the manuscript.

Funding

None.

Availability of data and materials

The source-code is available from the author on reasonable request. Gene4PD database is freely available at <http://genemed.tech/gene4pd>. Genotype-tissue expression (GTE_x) database is freely available at <https://gtexportal.org>.

Data availability

Data is available at <https://github.com/mmilano87/DNATool>.

Declarations

Competing interests

The authors declare that they have no Conflict of interest.

Received: 29 July 2024 Accepted: 24 October 2024

Published online: 18 November 2024

References

1. Guzzi PH, Cortese F, Mannino GC, Pedace E, Succurro E, Andreozzi F, Veltri P. Differential network analysis between sex of the genes related to comorbidities of type 2 mellitus diabetes. *Appl Netw Sci.* 2023;8(1):1–16.
2. Grimes T, Potter SS, Datta S. Integrating gene regulatory pathways into differential network analysis of gene expression data. *Sci Rep.* 2019;9(1):1–12.
3. Ideker T, Krogan NJ. Differential network biology. *Mol Syst Biol.* 2012;8(1):565.
4. Giancotti R, Vizza P, De Salazar M, Tradigo G, Guzzi PH, Irace C, Veltri P. Predicting glycemia by using rnns and heart rate patient data. In: 2023 International workshop on biomedical applications, technologies and sensors (BATS). IEEE; 2023, pp. 79–82

5. Gu S, Jiang M, Guzzi PH, Milenković T. Modeling multi-scale data via a network of networks. *Bioinformatics*. 2022;38(9):2544–53.
6. Chiarella G, Tognini S, Nacci A, Sieli R, Costante G, Petrolo C, Mancini V, Guzzi PH, Pasqualetti G, Cassandro E, et al. Vestibular disorders in euthyroid patients with hashimoto's thyroiditis: role of thyroid autoimmunity. *Clin Endocrinol*. 2014;81(4):600–5.
7. Succurro E, Marini MA, Fiorentino TV, Perticone M, Sciacqua A, Andreozzi F, Sesti G. Sex-specific differences in prevalence of nonalcoholic fatty liver disease in subjects with prediabetes and type 2 diabetes. *Diabetes Res Clin Pract*. 2022;190: 110027.
8. Guzzi PH, Cortese F, Mannino GC, Pedace E, Succurro E, Andreozzi F, Veltri P. Analysis of age-dependent gene-expression in human tissues for studying diabetes comorbidities. *Sci Rep*. 2023;13(1):10372.
9. Succurro E, Ciccone F, Papa A, Miceli S, Vizza P, Fiorentino TV, Perticone M, Sciacqua A, Guzzi PH, Veltri P, et al. Impaired insulin-stimulated myocardial glucose metabolic rate is associated with reduced estimated myocardial energetic efficiency in subjects with different degrees of glucose tolerance. *Cardiovasc Diabetol*. 2023;22(1):1–11.
10. Cannistraci CV, Valsecchi MG, Capua I. Age-sex population adjusted analysis of disease severity in epidemics as a tool to devise public health policies for covid-19. *Sci Rep*. 2021;11(1):1–8.
11. Tan BW, Tan BW, Tan AL, Schriver ER, Gutiérrez-Sacristán A, Das P, Yuan W, Hutch MR, Barrio NG, Jimenez MP, et al. Long-term kidney recovery and mortality after covid-19-associated acute kidney injury: an international multi-centre observational cohort study. *EClinicalMedicine*. 2023;55.
12. Mercatelli D, Pedace E, Veltri P, Giorgi FM, Guzzi PH. Exploiting the molecular basis of age and gender differences in outcomes of sars-cov-2 infections. *Comput Struct Biotechnol J*. 2021;19:4092–100.
13. Cannataro M, Guzzi PH, Mazza T, Tradigo G, Veltri P. Using ontologies for preprocessing and mining spectra data on the grid. *Future Gener Comput Syst*. 2007;23(1):55–60.
14. Mangoni M, Petrizzi F, Liorni N, Bianco SD, Biagini T, Napoli A, Adinolfi M, Guzzi PH, Novelli A, Caputo V, et al. Investigating mitochondrial gene expression patterns in drosophila melanogaster using network analysis to understand aging mechanisms. *Appl Sci*. 2023;13(12):7342.
15. Cho Y-R, Mina M, Lu Y, Kwon N, Guzzi PH. M-finder: uncovering functionally associated proteins from interactome data integrated with go annotations. *Proteome Sci*. 2013;11(1):1–12.
16. Milano M, Guzzi PH, Tymofieva O, Xu D, Hess C, Veltri P, Cannataro M. An extensive assessment of network alignment algorithms for comparison of brain connectomes. *BMC Bioinform*. 2017;18:31–45.
17. Roy A, Dunson DB. Nonparametric graphical model for counts. *J Mach Learn Res*. 2020;21(1):9353–73.
18. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. 2005;4(1).
19. Fuller TF, Ghazalpour A, Aten JE, Drake TA, Lusis AJ, Horvath S. Weighted gene coexpression network analysis strategies applied to mouse weight. *Mamm Genome*. 2007;18(6–7):463–72.
20. Stolovitzky G, Monroe D, Califano A. Dialogue on reverse-engineering assessment and methods: the dream of high-throughput pathway inference. *Ann NY Acad Sci*. 2015;1158(1):1–29.
21. Tesson BM, Breitling R, Jansen RC. Diffcoex: a simple and sensitive method to find differentially coexpressed gene modules. *BMC Bioinform*. 2010;11:497.
22. Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nat Rev Genet*. 2015;16(8):441–58.
23. Chuang H-Y, Lee E, Liu Y-T, Lee D, Ideker T. Network-based classification of breast cancer metastasis. *Mol Syst Biol*. 2007;3(1):140.
24. de la Fuente A. From 'differential expression' to 'differential networking'-identification of dysfunctional regulatory networks in diseases. *Trends Genet*. 2010;26(7):326–33.
25. Padi M, Quackenbush J. Differential network propagation: a new method to detect dysregulated pathways in complex diseases. *PLoS Comput Biol*. 2018;14(3):1006085.
26. Luo T, Zheng J. Network-based methods for the prediction of drug-target interactions using differentially expressed genes in disease. *Bioinformatics*. 2019;35(3):497–504.
27. Ha MJ, Baladandayuthapani V, Do K-A. Dingo: differential network analysis in genomics. *Bioinformatics*. 2015;31(21):3413–20.
28. Basha O, Argov CM, Artzy R, Zoabi Y, Hekselman I, Alfandari L, Chalifa-Caspi V, Yeger-Lotem E. Differential network analysis of multiple human tissue interactomes highlights tissue-selective processes and genetic disorder genes. *Bioinformatics*. 2020;36(9):2821–8.
29. Shojaie A. Differential network analysis: a statistical perspective. *Wiley Interdiscip Rev Comput Stat*. 2021;13(2):1508.
30. Lauritzen SL. Graphical models vol. 17. Clarendon Press, 1996.
31. Roy S, Mannes HN, Jha M, Guzzi PH, Kalita JK. Soft computing approaches to extract biologically significant gene network modules. *Soft Comput Biol Syst*. 2018, 23–37.
32. Galicia JC, Guzzi PH, Giorgi FM, Khan AA. Predicting the response of the dental pulp to sars-cov2 infection: a transcriptome-wide effect cross-analysis. *Genes Immun*. 2020;21(5):360–3.
33. Zimmerman DW. Comparative power of student t test and mann-whitney u test for unequal sample sizes and variances. *J Exp Educ*. 1987;55(3):171–4.
34. Allen GI, Liu Z. A local poisson graphical model for inferring networks from sequencing data. *IEEE Trans Nanobiosci*. 2013;12(3):189–98.
35. Besag J. Spatial interaction and the statistical analysis of lattice systems. *J R Stat Soc Ser B (Methodol)*. 1974;36(2):192–225.
36. Rani J, Mittal I, Pramanik A, Singh N, Dube N, Sharma S, Puniya BL, Raghunandan MV, Mobeen A, Ramachandran S. T2diacod: a gene atlas of type 2 diabetes mellitus associated complex disorders. *Sci Rep*. 2017;7(1):1–21.
37. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, Hasz R, Walters G, Garcia F, Young N, et al. The genotype-tissue expression (gtex) project. *Nat Genet*. 2013;45(6):580–5.
38. Pressler MP, Horvath A, Entcheva E. Sex-dependent transcription of cardiac electrophysiology and links to acetylation modifiers based on the gtex database. *Front Cardiovasc Med*. 2022, 9.

39. Ortuso F, Mercatelli D, Guzzi PH, Giorgi FM. Structural genetics of circulating variants affecting the sars-cov-2 spike/human ace2 complex. *J Biomol Struct Dyn*, 2021 , 1–11
40. Guzzi PH, Lomoio U, Veltri P. Gtexvisualizer: a web platform for supporting ageing studies. *Bioinformatics*. 2023;39(5):303.
41. Kanehisa M. The kegg database. In: 'In Silico'simulation of Biological Processes: Novartis Foundation Symposium 247, vol. 247. Wiley Online Library; 2002, pp. 91–103
42. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape stringapp: network analysis and visualization of proteomics data. *J Proteome Res*. 2018;18(2):623–32.
43. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. *Data mining in proteomics: from standards to applications*, 291–303 (2011)
44. Class CA, Ha MJ, Baladandayuthapani V, Do K-A. idingo-integrative differential network analysis in genomics with shiny application. *Bioinformatics*. 2018;34(7):1243–5.
45. Monaco C, Terrando N, Midwood KS. Toll-like receptor signaling: common pathways that drive cardiovascular disease and rheumatoid arthritis. *Arthritis Care Res*. 2011;63(4):500–11.
46. Yang Y, Lv J, Jiang S, Ma Z, Wang D, Hu W, Deng C, Fan C, Di S, Sun Y, et al. The emerging role of toll-like receptor 4 in myocardial inflammation. *Cell Death Dis*. 2016;7(5):2234–2234.
47. Abenavoli L, Luigiano C, Guzzi P, Milic N, Morace C, Stelitano L, Consolo P, Miraglia S, Fagoonee S, Virgilio C, et al. Serum adipokine levels in overweight patients and their relationship with non-alcoholic fatty liver disease. *Panminerva Med*. 2014;56(2):189–93.
48. Balistreri CR, Ruvolo G, Lio D, Madonna R. Toll-like receptor-4 signaling pathway in aorta aging and diseases: "its double nature". *J Mol Cell Cardiol*. 2017;110:38–53.
49. Tang G, Duan F, Li W, Wang Y, Zeng C, Hu J, Li H, Zhang X, Chen Y, Tan H. Metformin inhibited nod-like receptor protein 3 inflammasomes activation and suppressed diabetes-accelerated atherosclerosis in apoe^{-/-} mice. *Biomed Pharmacother*. 2019;119: 109410.
50. Al-Lamki RS, Mayadas TN. Tnf receptors: signaling pathways and contribution to renal dysfunction. *Kidney Int*. 2015;87(2):281–96.
51. Chen G, Goeddel DV. Tnf- α signaling: a beautiful pathway. *Science*. 2002;296(5573):1634–5.
52. Wang J, Knaut H. Chemokine signaling in development and disease. *Development*. 2014;141(22):4199–205.
53. Li S, Brown MS, Goldstein JL. Bifurcation of insulin signaling pathway in rat liver: mtorc1 required for stimulation of lipogenesis, but not inhibition of gluconeogenesis. *Proc Natl Acad Sci*. 2010;107(8):3441–6.
54. Noorolayi S, Shajari N, Baghbani E, Sadreddini S, Baradaran B. The relation between pi3k/akt signalling pathway and cancer. *Gene*. 2019;698:120–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.