Metabolomics (2026) 22:152
https://doi.org/10.1007/s11306-026-02529-y

ORIGINAL ARTICLE

# Network propagation in bipartite metabolite–reaction graphs for metabolomic data exploration

**Julia Kuligowski[1,2] · Marta Moreno-Torres[3,4,5] · David Pérez-Guaita[6] · Francesc Albert Esteve-Turrillas[6] · Guillermo Quintás[4,7]**

Received: 23 June 2026 / Accepted: 5 September 2026 / Published online: 16 September 2026
© The Author(s) 2026

## Abstract

**Introduction** Interpretation of metabolomic data is frequently limited by incomplete metabolite coverage and the predefined representation of biochemical organization provided by pathway-based approaches.

**Objectives** This study describes and evaluates a graph-based network propagation method designed to faciitate metabolomic data exploration by diffusing node-level statistical relevance scores across a metabolic network.

**Methods** An undirected bipartite metabolite–reaction network was reconstructed using the KEGG database, comprising 8236 nodes (1601 metabolites and 6635 reactions). Simulated datasets with 75% sparsity mimicking two metabolic perturbations and two no-effect control groups, and real data from a previous study were used to test the strategy. Node-level statistical signal scores were distributed across the network topology using a random walk-based diffusion operator combined with a supervised clamping procedure to preserve initially observed measurements. A topological distance mask was subsequently applied to restrict propagation to nodes located within a predefined network distance from experimentally measured metabolites.

**Results** Network propagation redistributed statistical relevance across connected subgraphs, expanding the number of metabolic features carrying topology-informed statistical scores beyond experimentally observed metabolites in simulated and real data. In both oxidative stress and mitochondrial dysfunction simulations, significant nodes displayed non-random topological organization consistent with the simulated perturbations. Propagation of real data also identified additional unmeasured metabolites that were topologically connected to experimentally observed metabolites. In multivariate analyses, network propagation expanded the feature space and showed that it could improve clustering performance.

**Conclusions** The proposed bipartite metabolite–reaction network propagation provides an exploratory tool for metabolomics. By integrating topological context with statistical relevance, this approach complements established pathway analyses for hypothesis generation and candidate feature recovery in partially observed metabolic systems.

**Keywords** Metabolomics · Multivariate analysis · Data exploration · Network propagation · Partial least squares – discriminant analysis (PLS-DA)

✉ Guillermo Quintás
gquintas@leitat.org

[1] Spanish Network in Maternal, Neonatal, Child and Developmental Health Research (RICORS- SAMID) (RD24/0013/0014), Neonatal Research Group, Health Research Institute La Fe, Valencia, Spain

[2] Servicio de Análisis de Vesículas Extracelulares (SAVE), Health Research Institute Hospital La Fe, Valencia, Spain

[3] Department of Biochemistry and Molecular Biology, University of Valencia, Valencia, Spain

[4] CIBER of Hepatic and Digestive Diseases, CIBER-EHD, 28029 Madrid, Spain

[5] Mixed Unit of Experimental Hepatology, Health Research Institute Hospital La Fe, Valencia, Spain

[6] Department of Analytical Chemistry, University of Valencia, Burjassot, Spain

[7] Health and biomedicine, Leitat Technological Center, Barcelona, Spain

## 1 Introduction

Multivariate statistical modeling has become a central tool in metabolomics for capturing coordinated patterns of metabolic variation associated with physiological states or external interventions. However, translating statistical model outputs into interpretable biological mechanisms remains challenging. While machine learning approaches often provide strong predictive performance, they typically offer limited mechanistic insight into the biological processes underlying model outputs. This has driven increasing interest in explainable artificial intelligence (XAI) methods (Ivanovs et al., 2021; Sidak et al., 2022), including perturbation-based strategies, which aim to identify relevant variables or groups of variables by assessing the impact of controlled input modifications on model predictions.

Interpretation of metabolomic data is commonly facilitated by grouping metabolites into biologically meaningful sets, most often defined by metabolic pathways. This approach has proven very useful for linking multivariate model behavior to known biochemical functions (Kuligowski et al., 2024). However, pathway-based representations present several limitations. Pathways are predefined, often inconsistently curated across databases, and do not fully capture the interconnection and overlap of metabolic processes (Green and Karp, 2006; Radrich et al., 2010). Moreover, metabolites frequently participate in multiple pathways, complicating the attribution of functional relevance, and metabolomic coverage is incomplete, thus limiting the interpretability of pathway analyses (Lee et al., 2025).

Network-based representations provide an alternative in which relationships between biochemical entities are represented. Previous approaches have typically used metabolite–metabolite networks, in which edges represent inferred or predefined relationships between metabolites (Kuligowski et al., 2025). However, metabolism is fundamentally mediated by enzyme-catalyzed reactions, and a reaction may involve multiple substrates and products. Representing metabolic systems as bipartite networks, where metabolites and reactions are explicitly modeled as distinct node types, preserves the relationships between metabolites and reactions, and the reaction dependencies. This representation additionally could allow statistical information derived from metabolomic measurements to be propagated to reaction nodes and subsequently to other metabolites connected through the reaction network.

Due to technical constraints in detection and annotation, only a subset of metabolites present in the biological system can be detected and a fraction of them annotated in metabolomic studies. Consequently, the experimentally observed metabolome represents a sparse projection of the underlying metabolic network. This sparsity can bias downstream analyses, particularly when estimating the contribution of specific network structures, whose relevance may be underestimated, as network regions close to measured metabolites may contain biologically relevant information that is not directly represented in the experimental feature matrix.

Network propagation refers to a class of graph-based methods that diffuse node-associated information across a network according to its connectivity structure (Cowen et al., 2017). Signals assigned to individual nodes are distributed to neighboring nodes, enabling the integration of local measurements with global network topology. Thus, network propagation (Pirhaji et al., 2016; Hui et al., 2023) enables the integration of biological signals across interaction networks, facilitating the characterization of biological mechanisms and providing insight into biological processes (Cowen et al., 2017). Over the last years, a number of bioinformatic techniques have been proposed for genetic, proteomic and other omic analysis using networks, based on random walks, information diffusion, and electrical resistance among others. These strategies have been successfully applied to identify disease genes, genetic modules and drug targets (Visonà et al., 2024). However, their application to metabolomic data, particularly using bipartite metabolite–reaction representations, remains limited.

In this work, we adapt network propagation to a bipartite metabolite-reaction network for exploratory metabolomic data analysis. In this approach, the signal associated with each metabolic node does not represent metabolite concentration or abundance, but a node level statistical score (SC), representing the strength of statistical evidence associated with a metabolite, derived from univariate or multivariate analyses. Thus, propagated SCs values represent topology-informed statistical relevance rather than reconstructed metabolic measurements. The approach is evaluated using controlled simulations representing two distinct metabolic perturbations and two no-effect scenarios, followed by application to an experimental human metabolomic dataset. We specifically assessed whether propagation extends statistical signal into topologically connected network regions, whether the resulting significant nodes exhibit non-random network organization, and whether propagation improves multivariate discrimination. Results show that by leveraging graph connectivity, nodes topologically adjacent to statistically significant metabolites accumulate signal, allowing the prioritization of unmeasured metabolites and reaction

nodes, providing a flexible, topology-guided approach for hypothesis generation and candidate feature identification.

## 2 Materials and methods

### 2.1 Bipartite reaction-metabolite network

An undirected bipartite reaction-metabolite network reconstruction was performed using the KEGG REST API (https://www.kegg.jp/kegg/rest/) to retrieve compound–reaction associations and reaction-level annotations. Metabolites and reactions were represented as distinct node classes, and edges connected metabolites to reactions in which they participate. KEGG compound identifiers were queried in batch mode, associated reactions were extracted from compound entries, and complete reaction equations were then retrieved.

To focus the analysis on core metabolic processes, metabolites were filtered to retain 1813 participating in at least one of a curated set of 53 metabolic pathways summarized in Table SM1, while excluding others such as drug metabolism, xenobiotic degradation, and other peripheral pathway classes. Additional filtering steps were applied to remove currency metabolites (e.g., ATP, NADH, $H_2O$) known to introduce non-specific connectivity. Generic KEGG compound entries representing generalized, specific classes of compounds were also excluded. The resulting network was exported for topological analysis in Gephi (v0.10.1). The network resulted in a graph with 8236 nodes (1601 metabolites and 6635 reactions) and 11,867 edges and it was visualized using the Force Atlas 2 layout. Community structure and topological partitioning were evaluated using the Louvain modularity optimization algorithm implemented in Gephi 0.10.1, using default edge weighting and a resolution parameter of 1.0.

This KEGG-derived biochemical network is not intended as an organism-specific metabolic reconstruction. It includes metabolites participating in at least one predefined set of KEGG metabolic pathways. Therefore, the inclusion of a metabolite or pathway in the network does not imply that the corresponding pathway is present or active.

### 2.2 Network signal propagation

To enable topology-guided diffusion across the graph, input data must be structured into a standardized feature matrix derived from single-study or multi-experiment workflows. The required input structure and algorithmic propagation steps are organized as follows: first, the initial feature matrix $\mathbf{X}$ ($n$ x $j_{exp}$) is set up, where $n$ represents the number of independent studies, bootstrapped iterations, or experimental comparisons, and $j_{exp}$ corresponds to the number of metabolites experimentally analyzed. Entries in $\mathbf{X}$ do not represent raw abundance or concentration levels. Instead, each vector $\boldsymbol{x_i}$ (1 x $j_{exp}$) contains node-level statistical significance score values (e.g. SC = -log(p-values)) from uni- or multivariate differential testing or regression models. The experimentally measured metabolites are mapped onto the network containing $j$ metabolites and $k$ reaction nodes. Measured metabolites retain their observed SC values, whereas unmeasured metabolites and reaction nodes, are initialized with SC = 0, generating the initial network matrix $\mathbf{X_{initial}}$ ($n$ x ($j+k$)).

To incorporate biochemical network structure into the analysis, node-level metabolite signals are propagated across the bipartite metabolite–reaction network (Frainay and Jourdan, 2017) using a graph diffusion framework (Zhou et al., 2003). Let $\mathbf{A}$ denote the adjacency matrix of the bipartite network, and $\mathbf{D}$ represent the diagonal matrix of node out-degrees. The transition probability matrix $\mathbf{W}$ is constructed via row-normalization as $\mathbf{W} = \mathbf{D}^{-1}\mathbf{A}$. Signal propagation is then performed using the closed-form diffusion operator $\boldsymbol{F} = (\boldsymbol{I} - \alpha\, \boldsymbol{W})^{-1}\boldsymbol{X_{initial}}$ , where $\mathbf{F}$ is the propagated signal matrix, $\mathbf{I}$ is the identity matrix, and α is the diffusion parameter ranging between 0 and 1. Lower α values limit diffusion to immediate topological neighbors, whereas higher values allow information to spread across broader network neighborhoods. In this work, α = 0.5 was predefined as a fixed intermediate value and applied unchanged to both simulated and experimental datasets.

Following propagation, a supervised clamping procedure was applied to preserve experimentally observed information. Metabolite nodes with initial SC > 0 were reset to their original values, while initially unmeasured metabolite and reaction nodes retained their propagated values.

Finally, a topological distance mask was applied to eliminate global background diffusion noise and restrict inferred signals. Shortest path distances from all network nodes to the nearest experimentally measured metabolite were computed across the network graph topology. Nodes exceeding a predefined distance threshold of $d_{max}$ steps (e.g. $d_{max}$=2) were considered topologically remote and had their propagated SC set to zero. The resulting matrix $\mathbf{X_{propagated}}$ ($n$ x $j+k$), was used as the expanded dataset for downstream unsupervised and supervised univariate and multivariate exploration.

An example of a small bipartite network consisting of 8 metabolite nodes and 10 reaction nodes is shown in Fig. 1. The figure shows the impact of different α values on the trade-off between local and global information flow, and how reaction node values are estimated as topology-aware aggregates of signals from both directly connected and indirectly related metabolites. Initially, the network contains only two non-zero metabolite nodes (C1 and C7),

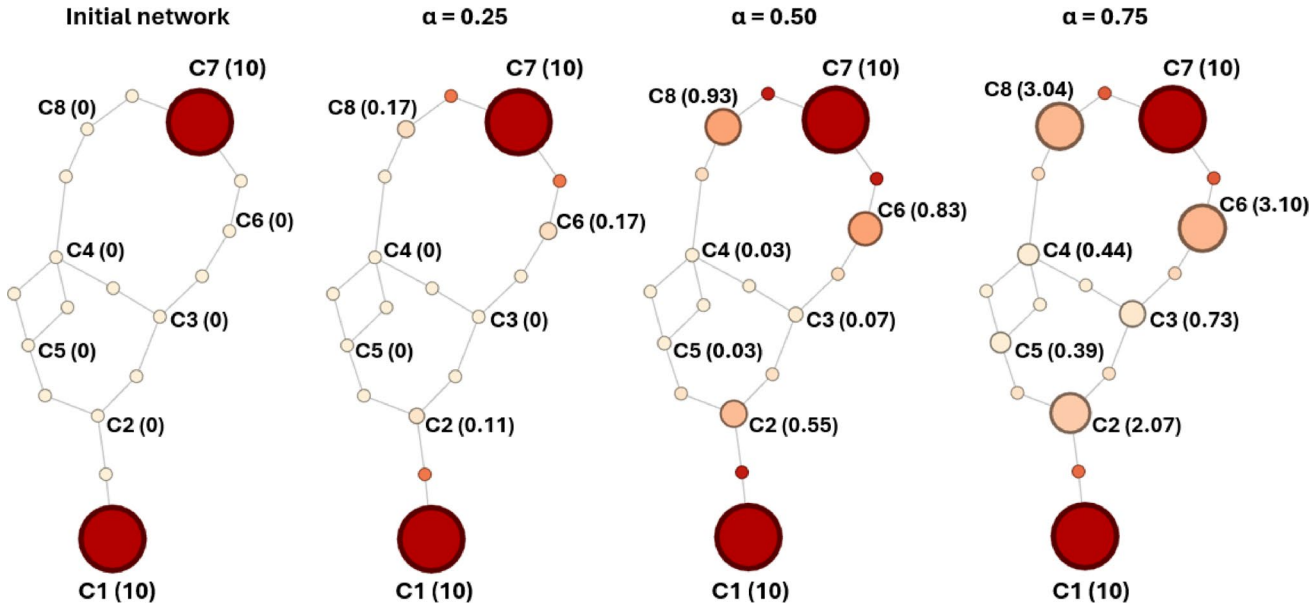


**Fig. 1** Effect of the parameter α on the network propagation of a small bipartite metabolite–reaction network. Circles connecting metabolites (C labeled nodes) represent reactions. The color and size of metabolic nodes represent the node SC. The size of reaction nodes was kept constant for better visualization

with SC = 10, respectively. Signal propagation was then performed using a diffusion parameter α = 0.25, 0.50 and 0.75. After propagation, a supervised clamping step restored the original values of C1 and C7. As α increased, the signal progressively diffuses from the initially non-zero nodes to neighboring reaction nodes and subsequently to more distant metabolites through shared connectivity. At low α (0.25), propagation remained localized, with signal change largely limited to immediate neighbors, and nodes C3, C4 and C5 remained unaffected. At intermediate α (0.5), signal spread more broadly while still preserving local structure, enabling nearby reactions to accumulate moderate propagated signal. At higher α (0.75), diffusion extends across larger portions of the network, leading to increased signal in distal nodes. In this work, a moderate value $\alpha = 0.5$ was selected to integrate network structure without excessive smoothing of the signal.

## 2.3 Simulated metabolomic data

Synthetic datasets were generated to represent two metabolic perturbation scenarios, i.e. oxidative stress (OS) and mitochondrial dysfunction (MD), and two no-effect scenarios (NE1 and NE2). OS represents a multi-pathway perturbation involving high-degree core hubs such as L-glutamate and glutathione. On the other hand, MD focused on the central energy-metabolism backbone and highly connected nodes (e.g., Acetyl-CoA and 2-oxoglutarate). The NE1 and NE2 groups were simulated using identical procedures but independent random seeds, ensuring that they share the same statistical properties while containing no systematic differences to enable the evaluation of null comparisons (NE1 vs. NE2). Each scenario (OS, MD, NE1, and NE2) simulated the outcome from 15 studies comparing control with a perturbed population. For example, in the MD scenario, each study (i.e., a row in $\mathbf{X}_{\mathbf{initial}}$ ($n$ x $j$)) could represent an experiment assessing the metabolomic effects of a different drug known to induce mitochondrial dysfunction in a biological system (Fig. 2).

A common set of 400 detected metabolites was used across all simulated datasets, corresponding to 25% of the 1601 metabolite nodes in the final network. The remaining metabolite nodes were assigned SC = 0, resulting in a shared missingness structure across conditions. This shared set included all metabolites included in Table 1.

Baseline values for the 'detected' (i.e. non-zero) metabolites were randomly sampled from a uniform distribution on [0, 1] and then -log transformed. Condition-specific metabolic profiles were then introduced by selectively perturbing the metabolites associated with OS or MD as summarized in Table 1. For these metabolites, SCs were reassigned by randomly sampling from a lower value uniform distribution ([0.015, 0.25]) and then $-\log_{10}$ transformed, generating subtle but structured differences relative to baseline. The sets of perturbed metabolites differed between OS and MD conditions, and no systematic perturbation was applied to any metabolite in the NE1 and NE2 vs. control metabolic profiles. This design ensured consistent metabolite detection across all samples while including biological differences through condition-specific perturbation of subsets

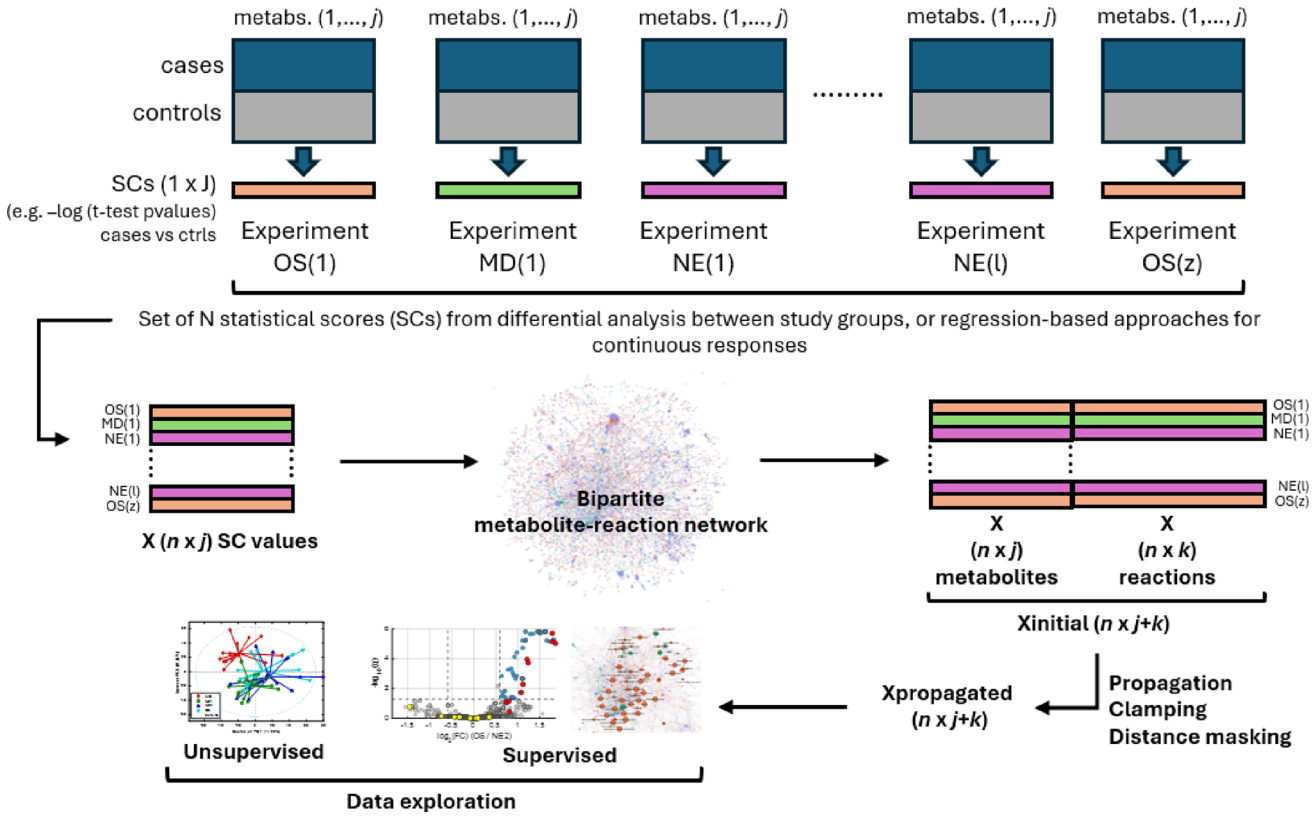


**Fig. 2** Scheme of the workflow for the analysis of the simulated data set. Results from a set of *n* independent studies comparing concentrations or levels of metabolites between two study groups, or regression-based approaches for continuous responses, are collected. For each study, a (1 x *j*) *statistical scores* (SCs) vector, representing the statistical significance of each metabolite in the network (e.g., t-tests -log(p-value)) are obtained. A SC = 0 is imputed for not detected metabolites. Then, the set of SC vectors is rearranged into a matrix $\mathbf{X}_{\mathbf{initial}}$ ($n$ x $j+k$) that includes the $k$ reaction nodes (SC = 0) and network propagation is carried out. The outcome $\mathbf{X}_{\mathbf{propagated}}$ ($n$ x $j+k$) provides an extended data matrix containing original and propagated values which is analyzed for data exploration

**Table 1** Sets of metabolites simulating perturbations associated with oxidative stress (OS) or mitochondrial dysfunction (MD)

| Effect | Systematically altered metabolites |
|---|---|
| oxidative stress (OS) | C00025 (Glutamic), C00037 (Glycine), C00051 (GSH), C00127 (GSSG), C00212 (Adenosine), C00242 (Guanine), C00294 (Inosine), C00387 (Guanosine), C01762 (Xanthosine), C01879 (Pyroglutamic) |
| mitochondrial dysfunction (MD) | C00022 (Pyruvic), C00024 (Acetyl-CoA), C00026 (Oxoglutaric), C00042 (Succinic), C00122 (Fumaric), C00149 (Malic), C00158 (Citric), C00186 (Lactic), C00424 (Lactaldehyde) |

of metabolites participating in selected mechanisms. As a result, inter-condition variation was driven by structured signal differences, and intra-condition variability arose from random intensity sampling. Reaction nodes were initialized to zero prior to propagation. Figure 3 shows the distribution of signals for boosted metabolites across conditions. Results depicted show that the effects in the metabolic profiles led to statistically significant differences relative to control (NE2) (two-sample t-test, $p < 0.05$) in a subset of the metabolites in each condition.

## 2.4 Experimental metabolomic dataset

The network-propagation method was evaluated on serum metabolic profiles from 1336 postmenopausal women enrolled in the Cancer Prevention Study II Nutrition Cohort were analyzed. The cohort is described in detail elsewhere(Stevens et al. 2018). The cohort comprised 332 women reporting estrogen-only use (E-only), 337 women reporting combined estrogen and progesterone use (E + P), and 667 women reporting no female hormone use (no use). Among the set of 781 metabolites identified included in the study, a subset of 229 were mapped to the reconstructed metabolic network. Non detected values were imputed as 1/3 of the minimum observed intensity of each metabolite.

The data set of intensities was used to generate a secondary data set by bootstrapping for the testing of the network propagation method. Two pairwise treatment-vs-reference comparisons were considered independently: E-only vs. no use and E + P vs. no use. For each comparison, 100 bootstrap resamples were generated by resampling subjects within the corresponding treatment and reference groups. For every

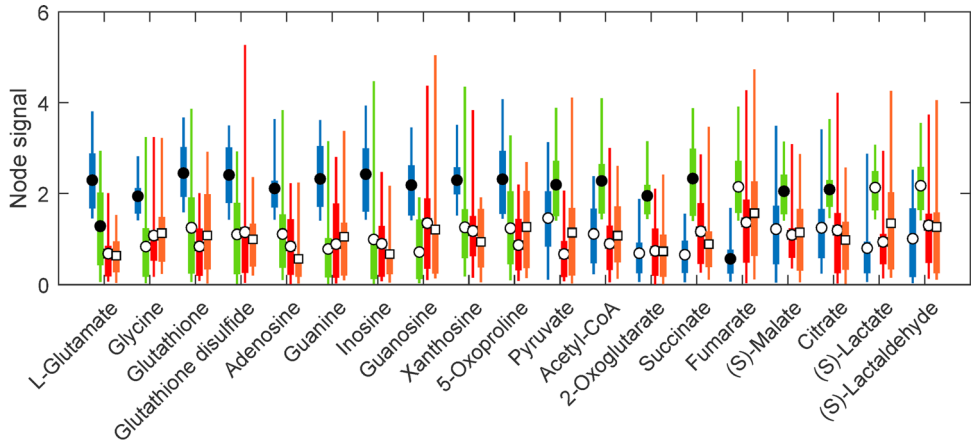


**Fig. 3** Simulated data set. Distribution of SCs for boosted metabolites across conditions. SCs obtained from 15 'case vs control' studies per condition are shown for oxidative stress (OS, blue), mitochondrial dysfunction (MD, green), and 'no effect' (NE1 (red) and NE2 (orange)). Vertical lines represent the 5th–95th percentile range, and thick bars indicate the interquartile range (25th −75th percentiles). Circles denote the mean SC value. Filled black circles indicate a statistically significant difference relative to NE2 (two-sample t-test, $p<0.05$), while open circles indicate non-significant differences (two-sample t-test, $p\geq0.05$)

bootstrap replicate, feature-wise statistical differences were calculated using Welch's two-sample t-test and transformed into statistical relevance scores (SCs) as SC original = -log10(p).

Measured metabolites retained their calculated SC values, whereas unmeasured metabolic nodes and the reaction nodes were initialized as SC = 0. After that, the 100 SC vectors obtained from the E-only vs. no-use and the 100 vectors obtained from the E + P vs. no-use were combined into a $\mathbf{X_{initial}}$ (200 × 8236) matrix for comparative analysis.

Network propagation was applied on $\mathbf{X_{initial}}$ using α = 0.5 and $d_{max}$=2, yielding $\mathbf{X_{propagated}}$ (200 × 8236). The resulting profiles represent repeated bootstrap estimates of the two treatment-versus-reference statistical signatures and not individual subject-level metabolomic observations. The two sets of bootstrap profiles were therefore treated as two classes of contrast specific statistical representations. The analysis was designed to identify network metabolites whose statistical relevance SC differed systematically between the E-only and E + P treatment contrasts. Metabolites that were not directly measured in the experimental dataset were considered network-inferred features and were interpreted as candidate extensions of the observed statistical signal.

## 2.5 Univariant discriminant analysis

For intensity data, log-transformed abundances were compared via parametric Welch's *t*-tests with Benjamini–Hochberg False Discovery Rate (FDR) correction (q).

For topology-propagated matrices of SC values, differential network node relevance across conditions was evaluated via empirical label permutation testing ($n_{perm}$ = 250). Empirical p-values were computed relative to the null distribution of absolute log-ratio differences across random class label assignments. Condition-specific topology differences were quantified using the statistical score ratio ( SCR), calculated as the ratio of mean propagated node scores between experimental groups (e.g., $SCR = \overline{SC}_{E+P} / \overline{SC}_{E-only}$). The SCR reflects relative changes in topology-propagated statistical relevance across the network and not absolute concentration or intensity fold-changes. The baseline mean score represents the absolute network relevance score, requiring a node to exhibit a minimum mean score of 0.5 in at least one experimental condition $max(\overline{SC}_{E+P}, \overline{SC}_{E-only})>0.5$). Significant network features were identified using multiple thresholding for p-value, SCR, and baseline mean SC to remove low-magnitude diffusion noise and non-informative background nodes from being selected as significant.

## 2.6 PLSDA

PLS-DA models were built on $\mathbf{X_{initial}}$ and $\mathbf{X_{propagated}}$ matrices to evaluate group separation. Model cross-validation was conducted using repeated stratified k-fold CV (5 repetitions, 5 folds) with autoscaling as preprocessing. Identical, balanced CV partitions were used for the initial and propagated models within each repetition, to ensure that differences in predictive performance were attributable to network propagation and not to variation in fold assignment. Latent variable selection was optimized to minimize CV classification error, and predictive performance was estimated using discriminant $Q^2$ ($dQ^2$) (Westerhuis et al., 2008).

Propagation-induced performance changes were calculated as $\Delta\, dQ^2 = dQ^2_{propagated} - dQ^2_{initial}$, where

$dQ^2_{propagated}$ and $dQ^2_{initial}$ were estimated using the propagated and initial networks, respectively. Statistical significance of $dQ^2$ and $\Delta dQ^2$ was evaluated using 250 label permutation iterations. Non-parametric one-sided p-values were estimated by comparing the observed statistics obtained using the true labels against their corresponding null distributions from permutation testing as:

$$p = \frac{\sum_{j=1}^{N_{perm}} 1\left(dQ^2_{null,j} \geq dQ^2_{observed}\right) + 1}{N_{perm} + 1}$$

### 2.7 Software and data availability

Simulations and statistical analysis were carried out in MATLAB 2023b (MathWorks) using custom scripts and the PLSToolbox 9.5.1 (Eigenvector Research Inc.). MATLAB scripts for calculating the standardized within-class sum of squares (Cabanski et al., 2010) (SWISS) were obtained from http://cancer.med.unc.edu/nhayes/pubs.html. Network analysis and visualization was carried out in Gephi (v0.10.1) running on Windows 11 (Microsoft). Over representation analysis (ORA)(Wieder et al., 2021) was carried out in Metaboanalyst 6.0(Pang et al. 2024) using hypergeometric tests against the KEGG library pathway. The experimental metabolomic data for this study was previously reported and described by Stevens et al. (2018) and it is deposited in the MetaboLights database with the identifier MTBLS136. Raw data and MATLAB code can be accessed on Zenodo (https://zenodo.org/records/20808246).

## 3 Results and discussion

### 3.1 Bipartite reaction-metabolite network

Figure 4 shows the resulting undirected bipartite metabolite–reaction network comprising 8236 nodes (1601 metabolic nodes) and 11,867 edges. The network exhibited a low average degree (2.88) and a large diameter (34), consistent with a bipartite topology in which connectivity describes biochemical dependencies rather than metabolite co-occurrence. The network was fragmented into multiple weakly connected components, consistent with pathway-level organization and partial disconnection across metabolic subsystems. Community detection was performed using the Louvain algorithm (Blondel et al., 2008) in Gephi 0.10.1 by maximizing modularity relative to a degree-preserving null model. The partition comprised 126 modules with a modularity value of 0.831, indicating a strong community structure across the network. The inclusion of reaction nodes facilitates the exploration of network topology by showing multi-substrate dependencies. This representation might also enable downstream assignment of reaction-level scores linked to metabolomic changes.

### 3.2 Analysis of simulated data

Unsupervised Principal Component Analysis (PCA) was used to characterize the global structure of the datasets before and after propagation (Fig. 5). The propagated representation was evaluated using all network nodes, metabolite nodes only, and reaction nodes only, while the original 400 detected metabolites provided the non-propagated reference.

To quantitatively evaluate the clustering of the different classes in the PC score space, the SWISS parameter was used (0 to 1 scale; lower values indicating stronger separation) and its statistical significance was estimated by permutation testing (Cabanski et al., 2010).

The network included 1601 and 6635 metabolic and reaction nodes, respectively. However, only 400 metabolites initially contained non-zero SC values. After propagation and topological distance masking, the number of non-zero metabolic and reaction network nodes increased up to 861 and 1863, respectively. Before propagation, the four simulated classes strongly overlap in the PC1-PC2 score space (SWISS = 0.624), in agreement with the simulation design in which perturbation effects were introduced as relatively small changes affecting a limited subset of metabolites.

Following network propagation, the overall distribution of samples in the PC space remained largely overlapping (Fig. 5B), but class organization changed, particularly when metabolite nodes were considered (Fig. 5C-D). The decrease in SWISS values indicates improved class separation, where OS and MD samples clustered more tightly along PC2, while separating from the overlapping no effect groups (NE1 and NE2) along PC1.

Univariate supervised analysis (Welch's test) and the calculation of the SCR were performed on the network-propagated data matrices for each node across three pairwise comparisons: (i) OS vs. NE2, (ii) MD vs. NE2, and (iii) a null comparison between two no-effect groups (NE1 vs. NE2). Features were classified as significant if q-value < 0.05 and $|\log_2(SCR)| > 0.5$. The volcano plots shown in Fig. 6 depict results obtained.

Network propagation assigned topology-informed SC values to metabolites that were not directly measured but were connected to experimentally observed regions of the network (Fig. 6). These network-inferred nodes should therefore be distinguished from experimentally measured metabolites because their propagated SCs represent statistical relevance inferred from network topology and not experimentally determined metabolite abundances. Consequently, the statistical differences observed for inferred

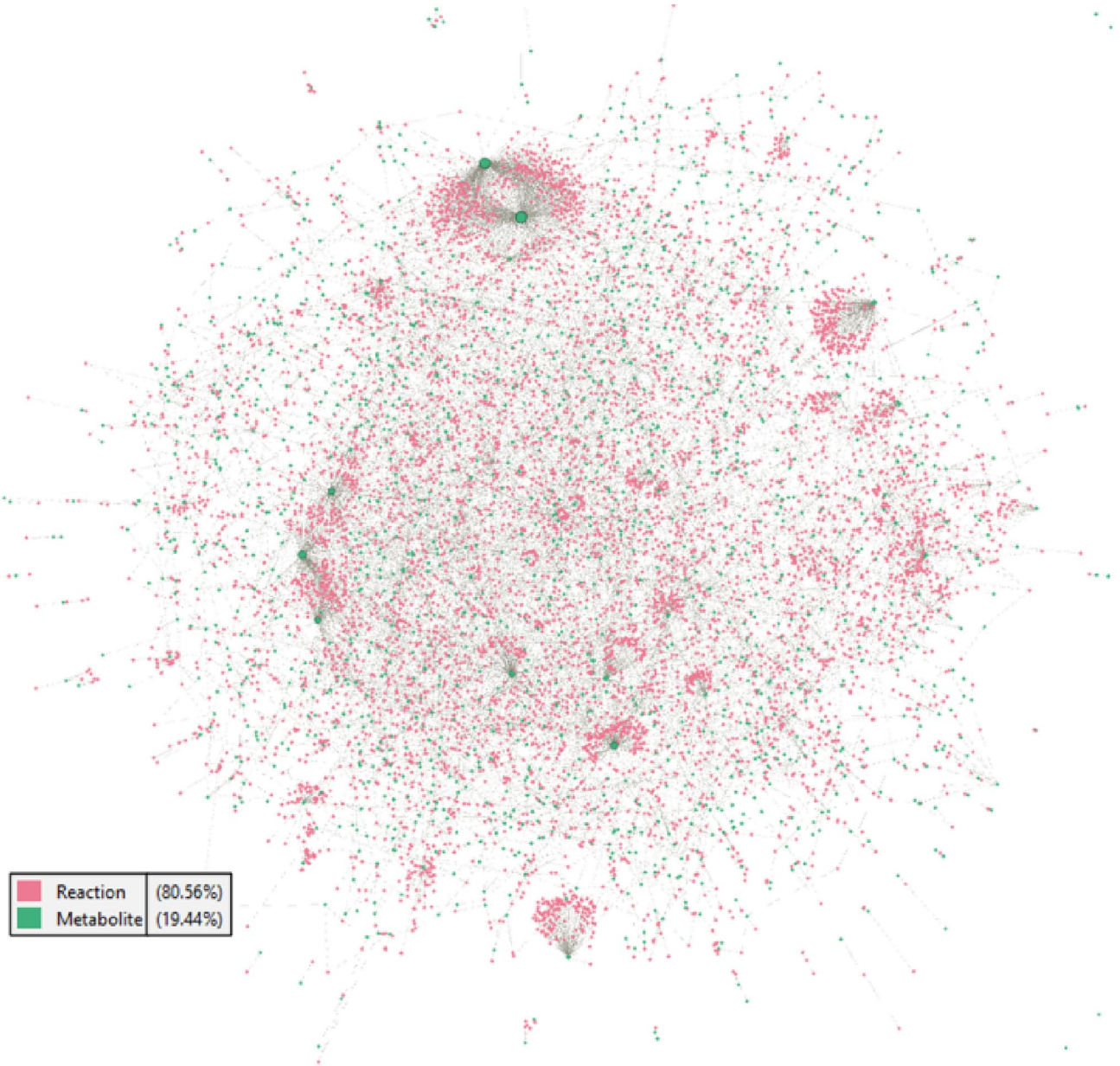


**Fig. 4** Structure of the metabolic network. Nodes represent metabolites (green) or reactions (purple). Edges represent directed associations through a chemical reaction. Node sizes are proportional to the node degree

nodes indicate differences in propagated relevance between conditions and should not be interpreted as direct evidence that the corresponding metabolites were differentially abundant. Instead, these features represent candidate extensions of the experimentally observed signal that may provide additional hypotheses for downstream investigation.

No metabolic or reaction nodes were found statistically significant in the null comparison between the no-effect groups (NE1 vs. NE2).

Figure 7 depicts a section of the network showing the names of the 99 metabolites found significant in the OS vs. NE2 analysis, including initially non-zero (green circles) and propagation-inferred (orange circles) metabolic nodes. The significant metabolites were concentrated in network regions associated with the simulated oxidative stress perturbation, particularly around glutathione metabolism, redox-related processes, and purine metabolism. The initially boosted metabolites included key components such as glutathione (GSH), glutathione disulfide (GSSG), L-glutamate, L-cysteine, and glycine. Following network propagation, additional topologically connected metabolites were identified, including glutathione-related compounds and intermediates such as γ-L-glutamyl-L-cysteine, 5-oxoproline, Cys-Gly, S-glutathionyl-L-cysteine, S-sulfanylglutathione,

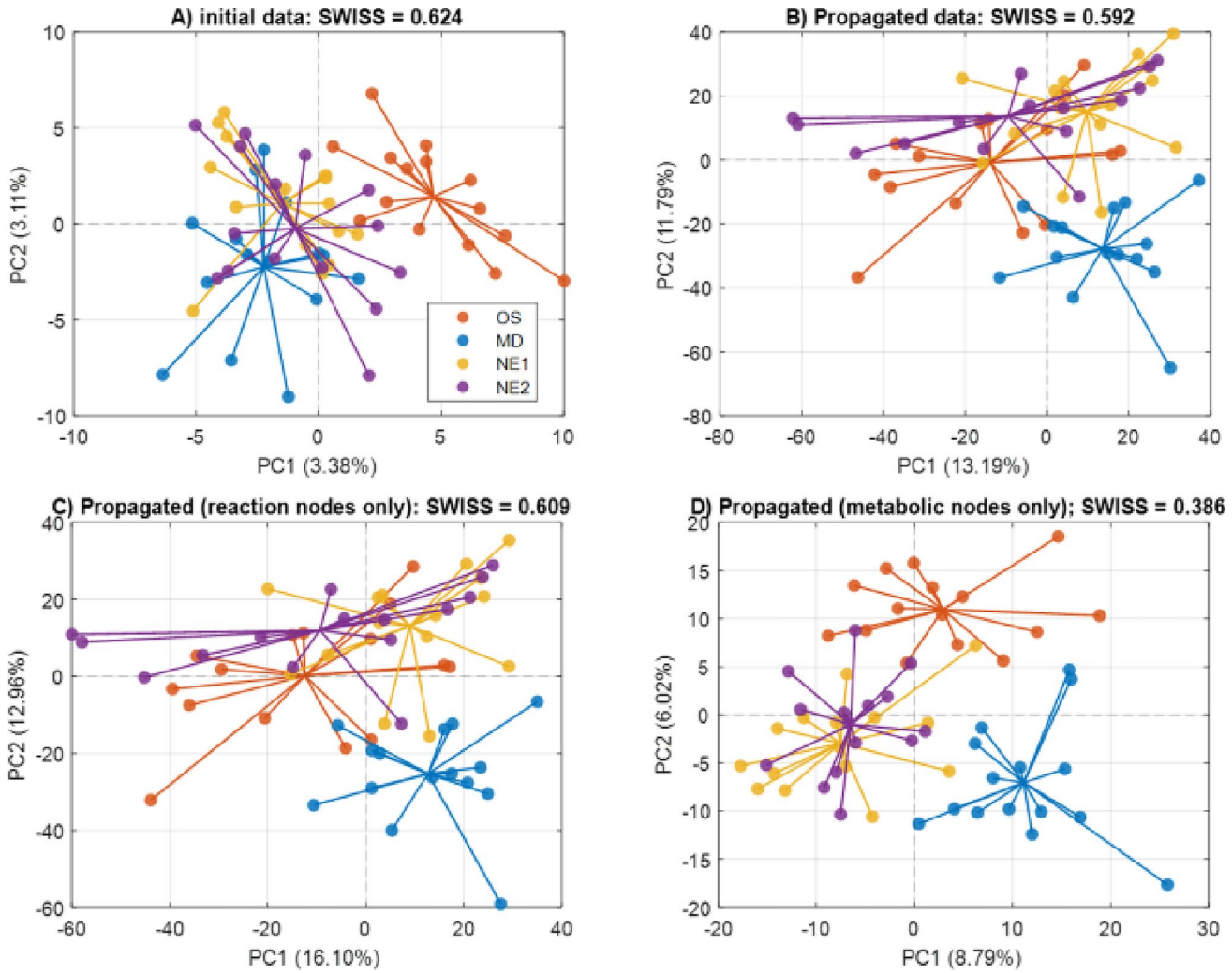


**Fig. 5** PCA scores plot of simulated network profiles. **A** original data (400 metabolite nodes); **B** propagated data including metabolite (861) and reaction (1863) non zero nodes; **C** propagated data including only reaction (1863) non zero nodes; **D** propagated data including only metabolite (861) non zero nodes

and glutathionylspermidine. Additional significant nodes were associated with purine and nucleotide metabolism, including IMP, GMP, adenine, adenosine, guanine, hypoxanthine, and inosine.

To evaluate whether the spatial distribution of significant metabolites reflected localized biological coordination, we implemented two complementary, degree-matched permutation tests ($n = 1000$). First, an empirical null distribution for the Largest Connected Component (LCC) was constructed by evaluating subgraphs induced strictly by reaction nodes bridging at least two target metabolites. Second, we calculated the Mean Shortest Path Distance ($d_{mean}$) across all unique pairs of target metabolites to assess overall spatial proximity across the network. For every permutation iteration, background node sets were selected to strictly match the exact degree distribution of the target metabolites.

The significant metabolites demonstrated highly non-random topological organization across both metrics. The LCC incorporated 98 of the 99 target metabolites into a single interconnected subnetwork ($p < 0.001$). Furthermore, the target metabolites exhibited a mean shortest path distance of 5.13 ($p < 0.001$) (Figure SM1). Together, the statistical significance of these metrics indicates that the perturbed metabolites form a non-random, spatially localized, and cohesive network neighborhood rather than a dispersed pattern across the network topology.

Figure 8 shows a section of the metabolic network showing metabolite names among the set of nodes found significant in the MD vs. NE2 analysis. In this simulation, the boosted metabolites targeted central carbon metabolism, including TCA cycle and intermediates such as pyruvate, citrate, 2-oxoglutarate, succinate, fumarate, malate, and oxaloacetate. Propagation extended the statistical signal

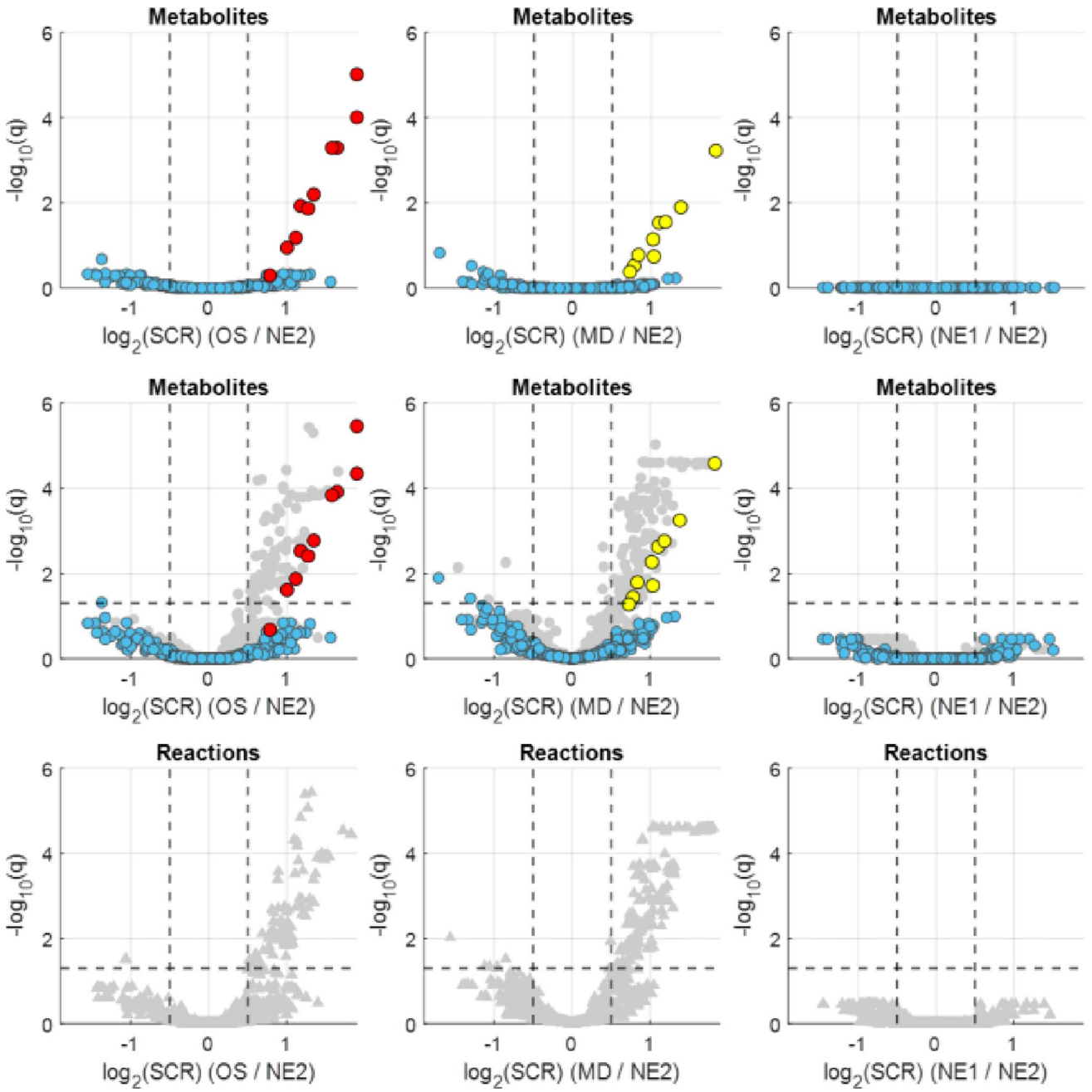


**Fig. 6** Volcano plots showing the $\log_2$(SCR) as a function of statistical significance ($-\log_{10}$ of False Discovery Rate-adjusted q-values) for each node, in the original (top row) and propagated networks (middle and bottom rows), in the comparisons OS vs. NE2 (left), MD vs. NE2 (center), and NE1 vs. NE2 (right) (top and middle rows: metabolites, bottom row: reactions). Metabolites are represented as circles (gray: not initially detected; blue: initially detected; red and yellow: initially detected and included in the sets of metabolites simulating perturbations associated with OS (red) or MD (yellow). Dashed lines indicate the thresholds for statistical significance ($q < 0.05$) and statistical relevance scores $|\log_2(SCR)| > 0.5$

from the initially perturbed central metabolic region into topologically connected biochemical pathways. Following network propagation, the set of 158 significant metabolites included also metabolites connected to these regions such as isocitrate, oxalosuccinate, acetoacetyl-CoA, succinyl-CoA, and multiple amino acid and CoA-associated intermediates. The propagated set also included metabolites associated with amino acid metabolism, such as L-alanine, L-asparagine, L-leucine, L-isoleucine, L-proline, and 2-oxoglutaramate, as well as intermediates involved in fatty-acid and branched-chain amino acid metabolism. The broad distribution of significant nodes is consistent with the central position and high connectivity of the boosted metabolites within the metabolic network. Using the degree-matched

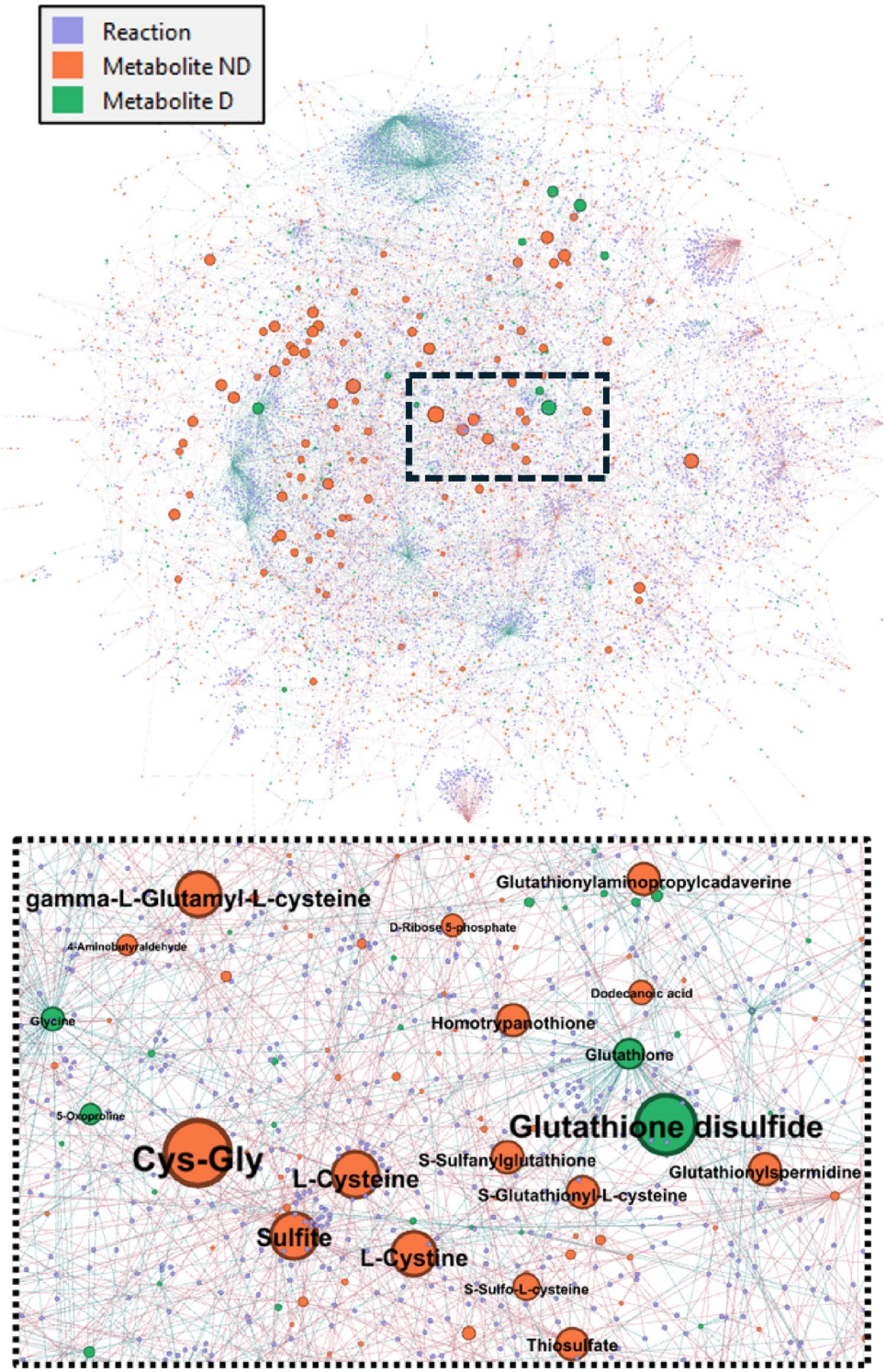


**Fig. 7** Top) OS vs. NE2 univariate analysis of the propagated simulated data set. Metabolic node size is proportional to the $-\log_{10}$ of FDR-adjusted q-values. Green nodes: simulated initially non-zero metabolites, orange nodes: propagation-inferred metabolites, purple nodes: reactions; Bottom) Section of the network showing names of some of the metabolites found significant in the OS vs. NE2 univariate analysis after propagation

permutation metrics (LCC and $d_{mean}$) topological testing described before confirmed that the set of 158 metabolites was non-randomly distributed across the network. The metabolite-centric LCC included 154 of 158 target nodes into a single subnetwork ($p<0.001$). Additionally, the significant metabolites exhibited mean shortest path distance of 5.03 ($p<0.001$) (Figure SM1).

Graph topology influences the distribution of the signal during network propagation. Highly connected or topologically central metabolites can receive or distribute signal

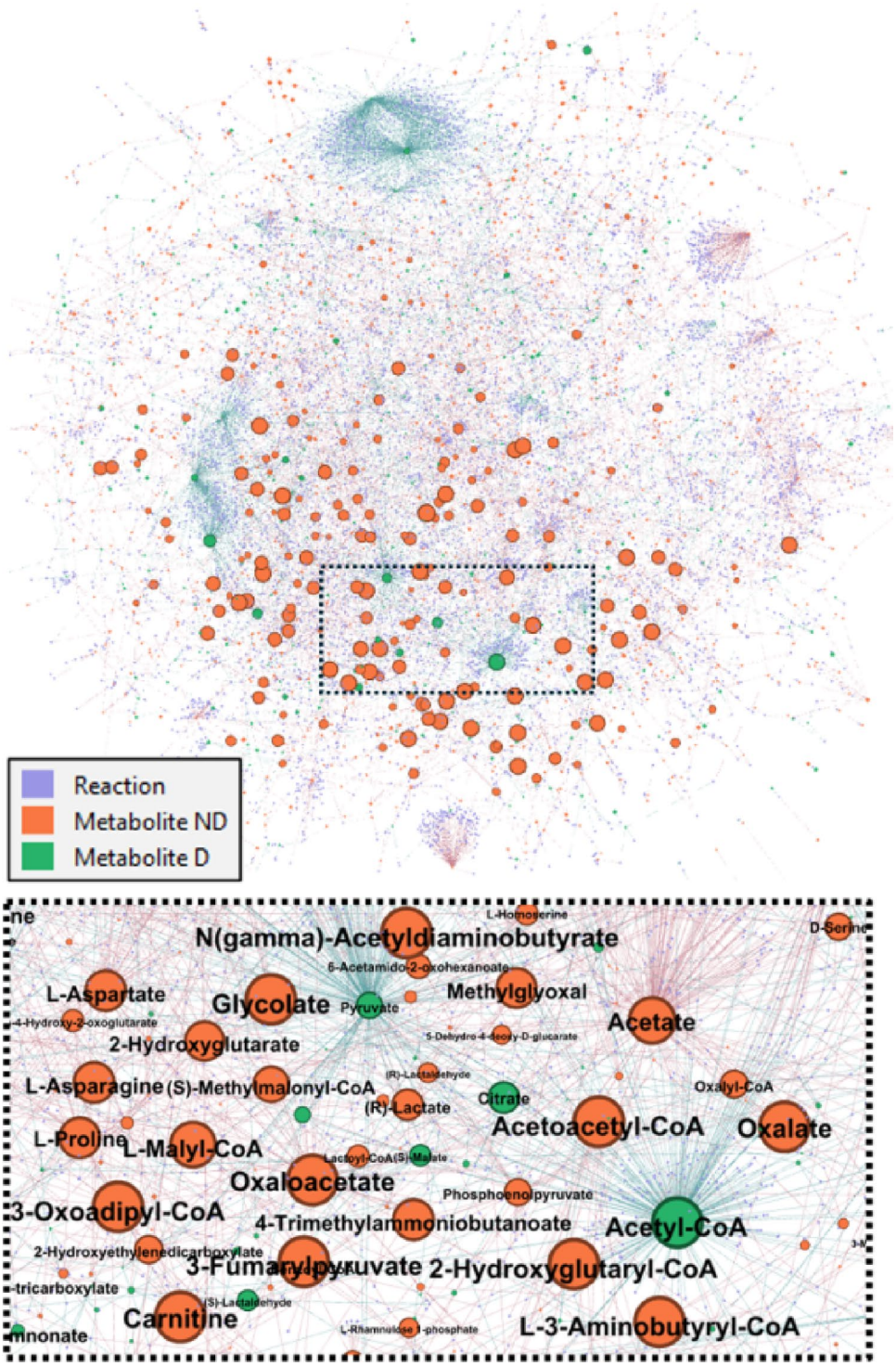


**Fig. 8** Top) MD vs. NE2 analysis of the propagated simulated data set. Metabolic node size is proportional to the $-\log_{10}$ of False Discovery Rate-adjusted q-values. Green nodes simulated initially non-zero metabolites, orange nodes: propagation-inferred metabolites, purple nodes: reactions; Bottom) Section of the network showing names of some of the metabolites found significant in the MD vs. NE2 NE2 univariate analysis after propagation

across a larger number of connected nodes due to their multiple connections within the network. In this simulation, the metabolites defining the OS and MD cases showed different topological properties. MD metabolites showed median betweenness and closeness percentiles of 99.9 (range 20–100) and 83.5 (range 64–85), respectively, indicating that these nodes occupy highly central and well-connected positions in the network. In contrast, selected OS metabolites also displayed high betweenness centrality (median 96.1, range 79.4–100), but lower closeness (median 64.6, range 37.4–84.2), suggesting a less uniformly central distribution. Several MD metabolites also corresponded to high-degree hubs, including 2-oxoglutarate (degree: 308), Acetyl-CoA (degree: 246), Pyruvate (degree: 211), and

Succinate (degree: 200)) (see Fig. 8). In contrast, OS metabolites had relatively low degree (< 10), with Glutamate being the exception (degree > 200). These topological differences were consistent with the broader propagation observed in the MD simulation compared with the more localized propagation pattern observed in OS.

To evaluate the impact of the simulated perturbations at the pathway level, over-representation analysis (ORA) was carried out using the KEGG metabolic pathway library. Enrichment-based approaches assess whether functionally related metabolites are overrepresented among discriminant metabolites, enabling pathway-level interpretation of coordinated metabolic responses. Consequently, only selected discriminant metabolites contribute to pathway enrichment. First, the simulated data set before propagation was analyzed. In the OS vs. NE2 analysis, ORA identified pathways associated with glutathione metabolism (Hits 3/28, FDR = 0.008), and purine metabolism (Hits 4/70, FDR = 0.008) as the most enriched metabolite sets. In the MD vs. NE2 analysis, ORA highlighted pathways related to the citrate cycle (TCA cycle) (Hits 4/20, FDR = $2\ 10^{-6}$), alanine, aspartate and glutamate metabolism (Hits 3/28, FDR = $9\ 10^{-4}$), and pyruvate metabolism (Hits 2/22, FDR = 0.02), as the most enriched metabolite sets.

ORA of the significant metabolites identified in the OS vs. NE2 comparison revealed five significantly enriched KEGG pathways after false discovery rate correction (FDR < 0.05) (Table SM2). The strongest enrichment was observed for purine metabolism, with 15/70 hits (FDR = $7\ 10^{-6}$). Alanine, aspartate and glutamate metabolism and glutathione metabolism were also strongly enriched, each with 9/28 hits (FDR = $3\ 10^{-5}$). Additional significant enrichment was observed for sulfur metabolism (4/8 hits, FDR = 0.004) and arginine biosynthesis (4/14 hits, FDR = 0.04). Together, these pathways highlight a concentration of the significant network propagated signal in purine metabolism and glutathione- and sulfur-related pathways, consistent with the predefined biochemical structure of the OS simulation. These results were consistent with the topological distribution of significant metabolites observed in Fig. 7, where metabolites associated with glutathione turnover and purine catabolism formed connected regions surrounding the initially perturbed nodes.

Pathway over-representation analysis of the 158 significant metabolites identified in the MD vs. NE2 comparison revealed 15 significantly enriched KEGG pathways (FDR < 0.05) (Table SM2). The strongest enrichment was observed for pathways directly related to central carbon metabolism, including the citrate cycle (TCA cycle; 12/20 hits, FDR = $10^{-9}$), pyruvate metabolism (12/22 hits, FDR = $2\ 10^{-9}$), and alanine, aspartate and glutamate metabolism (13/28 hits, FDR = $3\ 10^{-9}$). Additional significant enrichment was observed for glyoxylate and dicarboxylate metabolism (10/32 hits, FDR = $2\ 10^{-5}$), glycolysis/gluconeogenesis (5/24 hits, FDR = 0.04), and related short-chain carbon pathways including butanoate and propanoate metabolism. The enriched pathways also extended to amino-acid and lipid metabolism. Overall, the ORA profile was dominated by interconnected pathways involved in central carbon metabolism, amino-acid utilization, and fatty-acid metabolism, consistent with the broad network-level perturbation observed in the MD simulation.

In summary, the ORA profiles differed between the two simulated perturbations. Whereas the MD simulation produced broad enrichment across central carbon, amino-acid, and fatty-acid metabolic pathways, the OS simulation showed a more focused enrichment of purine, glutathione, sulfur, and related amino-acid pathways. This difference was consistent with the different network level distributions of the propagated signal observed for the two perturbation scenarios (Figs. 7 and 8). Nevertheless, the propagated feature set includes topology-inferred metabolites and these pathway enrichments should be interpreted as extensions of the statistical signal within the reference metabolic. Accordingly, the ORA results should also be considered exploratory and hypothesis-generating.

Propagation depends on both the diffusion parameter α and the maximum topological distance $d_{max}$. So, the sensitivity of the number of significant metabolites to α across a range of values and under two distance-masking conditions ($d_{max}$ = 0 and 2) was evaluated (Fig. 9). The effect of α differed between the two perturbation scenarios. For OS, the number of significant metabolites remained stable when $d_{max}$ = 2, whereas MD showed a stronger dependence on α, with the number of significant nodes increasing at higher diffusion values. Without distance masking, feature counts were substantially larger, particularly for MD, consistent with broader signal dispersion around highly connected seed nodes. These results demonstrate that propagation parameters influence the extent of inferred network signal and support the use of sensitivity analyses when selecting α and $d_{max}$.

Predictive modeling was performed using PLS-DA applied to both the non-propagated and propagated data networks, across the same three pairwise comparisons described above.

In the OS vs. NE2 comparison, the initial data set including 400 non-zero metabolic nodes showed statistically significant predictive performance ($dQ^2$ = 0.648, p-value = 0.008). After network propagation, performance increased to $dQ^2$ = 0.765 (p-value = 0.03), corresponding to a propagation-induced change in predictive performance $\Delta dQ^2$ = 0.117 (p-value = 0.04). In the MD vs. NE2 comparison, the initial model showed lower predictive performance ($dQ^2$ = 0.483,

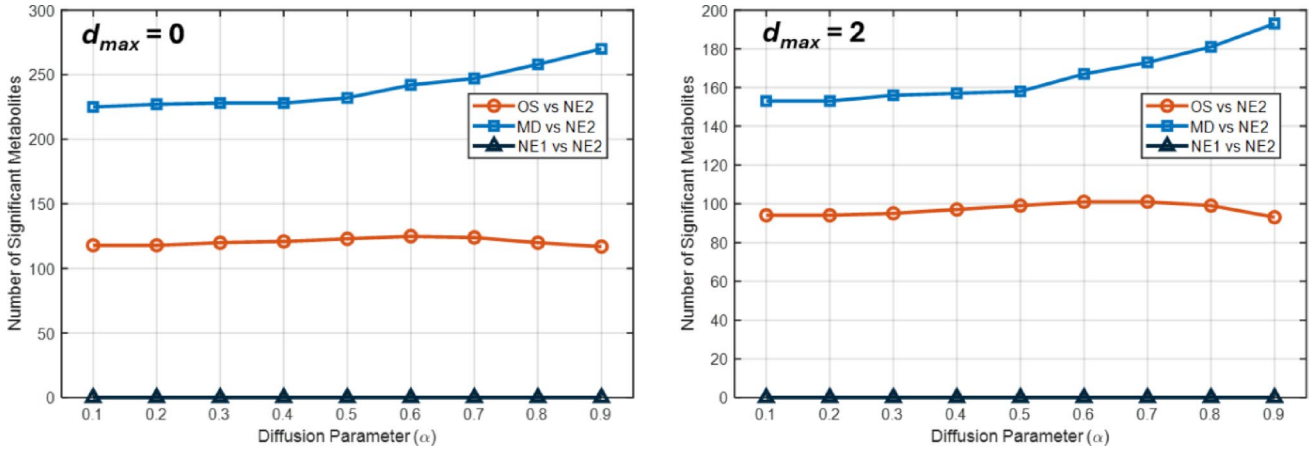


**Fig. 9** Effect of the diffusion parameter α and the maximum distance used in the topological distance masking ($d_{max}$) on the number of statistically significant metabolites identified, after network propagation in the OS vs. NE2, MD vs. NE2, and NE1 vs. NE2 comparisons. Significant metabolites were defined using q-value < 0.05 and $\log_2$SCR > 0.5 thresholds

p-value = 0.05) and after propagation, model performance improved ($dQ^2$ = 0.832, p-value = 0.01), showing a statistically significant positive propagation-induced change in predictive performance ($\Delta dQ^2$ = 0.349, p-value = 0.004). Thus, in these cases, propagation produced a statistically detectable improvement in predictive discrimination, although the magnitude of the improvement can be considered limited relative to the increase in feature dimensionality.

In the null comparison (NE1 vs. NE2), neither the initial ($dQ^2$ = 0.363 p-value = 0.9) nor propagated ($dQ^2$ = 0.293, p-value = 0.9) representations showed statistically significant class discrimination. The propagation-induced change in $dQ^2$ ($\Delta dQ^2 = -0.07$, p-value = 0.7) was also not significant. This indicates that propagation can redistribute background statistical variation across the network without generating artificial class-specific predictive signal.

These results indicate that propagation does not always improve predictive discrimination. In the no-effect comparison, neither the original nor propagated representation showed significant discrimination. Instead, its effect on predictive performance depends on the topology of the perturbed network regions and the relationship between those regions and the experimentally observed features.

### 3.3 Application to an experimental metabolomic dataset

To evaluate the applicability of the bipartite network propagation framework on biological data, serum metabolic profiles from 1336 postmenopausal women enrolled in the Cancer Prevention Study II Nutrition Cohort (MetaboLights ID: MTBLS136) were analyzed(Stevens et al. 2018). Cohort participants were categorized into three distinct postmenopausal hormone therapy groups: estrogen-only users ('E-only', $n$ = 332), combined estrogen and progesterone users ('E + P', n = 337), and non-users ('no use', n = 667). A total of 229 measured metabolites mapped directly onto the reconstructed KEGG reference network and were retained for downstream scoring. Feature-wise differential abundance was evaluated comparing the three groups pairwise (i.e. 'E-only' vs 'no use', 'E + P' vs 'no use', and 'E + P' vs 'E'). Then, $\mathbf{X_{initial}}$ (30 × 8236) was propagated as described before using a α = 0.5 and $d_{max}$ = 2.

Univariate differential analysis of metabolite intensities across the unpropagated hormone therapy cohorts revealed distinct baseline alterations. Fold-change ratios were calculated between treatment conditions. Estrogen only therapy (E-only vs. no use) showed widespread changes, yielding 19 significantly altered metabolites. These included altered steroid hormone derivatives (e.g., Estrone 3-sulfate, Androstenediol, and Dehydroepiandrosterone sulfate), functional fatty acids (including DPA and hexadecenoic acid), and non-steroidal pathways such as one-carbon/amino acid metabolism (Betaine, Sarcosine, L-Pipecolate, L-Cystathionine), and gut microbiome degradation products (Quinate, 4-Hydroxy-2-quinolinecarboxylic acid). Combined therapy (E + P vs. No use) significantly altered 6 metabolites, primary among them being the xenobiotics Quinate and Acetaminophen, the phospholipid intermediate Ethanolamine phosphate, and a sulfated steroid (Dehydroepiandrosterone sulfate). Direct comparison between treatments (E + P vs. E-only) identified 6 significant metabolites (including Androstenediol and Dehydroepiandrosterone sulfate) that differentiate both therapies.

For the topology-propagated SC matrices generated by bootstrap, differential network node relevance between hormone therapies was evaluated via empirical label permutation testing ($n_{perm}$ = 250). Condition-specific topology differences were quantified using the SCR, calculated as the

ratio of mean propagated node scores between experimental groups ($SCR = \bar{SC}_{E+P} / \bar{SC}_{E-only}$). Significant network features were identified using p-value<0.05, SCR<0.8 or >1.25, and $max(\bar{SC}_{E+P}, \bar{SC}_{E-only}) > 0.05$.

Univariate evaluation of the topology-propagated SC matrix identified a total of 51 significant metabolite features (Fig. 10) corresponding to 41 experimentally measured metabolites and 10 network-inferred metabolites. The directly measured metabolites included steroid metabolites (Androstenediol, Dehydroepiandrosterone sulfate, Cortisone, Deoxycholic acid), purine/pyrimidine intermediates (Urate, Pseudouridine, 3-Ureidopropionate, Deoxyuridine), lipid components (Sphinganine, Icosanoic acid, Arachidonate), and amino acid derivatives (N-Acetylaspartylglutamate, L-Cystine, S-Sulfo-L-cysteine).

The 10 network inferred metabolites obtained statistical relevance through propagation. Thus, these nodes should be interpreted as candidate metabolites associated with the observed network signal rather than as experimentally demonstrated differential metabolites. The list included 2-Hydroxyestrone (SCR = 0.32, p-value = 0.02), S-Adenosyl-L-methionine (SCR = 1.7, p-value = 0.02), 5-(5-Phospho-D-ribosylaminoformimino)−1-(5-phosphoribosyl)-imidazole-4-carboxamide (SCR = 2.5, p-value = 0.004), (S)-Acetoin (SCR = 2.5, p-value = 0.02), D-Ribose (SCR = 4.3, p-value = 0.01), 3-Dehydroquinate (SCR = 4.4, p-value = 0.02), BQ 123 (SCR = 4.8, p-value = 0.01), 16-Glucuronide-estriol (SCR = 5.0, p-value = 0.01), 4-Phospho-L-aspartate (SCR = 5.0, p-value = 0.01), and 2,5-Diamino-6-(5-phospho-D-ribosylamino)pyrimidin-4(3 H)-one (SCR = 5.5, p-value = 0.008).

This highlights the capability of the network diffusion model to highlight unmeasured metabolites such as estrogen phase-I/II metabolism, methyl donor pathways, and histidine/purine precursors that were undetected in standard feature-level metabolomic profiling. Specifically, inferring a phase II estrogen metabolite (16-Glucuronide-estriol) directly through topological connections to estrogenic metabolites and glucuronide donors demonstrate the utility of bipartite network diffusion for prioritizing unmeasured candidate biomarkers. Modularity analysis also showed that 16-Glucuronide-estriol, 2-Hydroxyestrone, D-Glucuronate, Androstenediol, Dehydroepiandrosterone sulfate, and Cortisone co-clustered within the same structural network community, reflecting clear functional modularity in steroid biotransformation pathways.

To assess global topological clustering across the reference network, we conducted degree-matched permutation testing ($n = 1000$). In contrast to focused simulated perturbations, the empirical target metabolites yielded a metabolite-centric LCC = 15/51 ($p = 0.5$), and a mean pairwise distance $d_{mean} = 8.13$ ($p = 0.3$) (Figure SM2). The lack of global significance under these metrics indicates that the experimental signal does not cluster into a single subgraph of close metabolites. Instead, the propagated statistical signal is distributed across multiple distinct, multi-focal network areas linked to metabolites involved if different pathways (e.g., steroid biotransformation, purine metabolism, lipid pathways) that form localized functional modules without global network consolidation. These results illustrate how network propagation can prioritize unmeasured metabolites that are locally connected to experimentally observed statistical signals within functional sub-pathways, providing structured candidates for targeted follow-up and validation.

PLS-DA of the bootstrap-derived contrast profiles showed significant discrimination between the E-only and E + P contrasts both before ($dQ^2 = 0.992$, mean $dQ_{null}{}^2 = 0.458$, p-value = 0.004) and after propagation ($dQ^2 = 0.984$, mean $dQ_{null}{}^2 = 0.526$, p-value = 0.004). However, propagation did not significantly change predictive performance ($\Delta dQ^2 = -0.007$, $\Delta dQ_{null}{}^2 = 0.07$, p-value = 0.9). Thus, in the experimental dataset, propagation substantially altered the network representation and enabled identification of additional topology-inferred features, but did not provide evidence of improved predictive discrimination.

Some cautions should be considered when interpreting network-propagated metabolomic data. First, network propagation processes redistribute signals across the topology and it does not provide independent experimental evidence for the inferred nodes receiving propagated signal. Inferred scores should consequently be interpreted as topology-informed statistical relevance. Second, the extent of propagation depends on the experimental data, network topology and on the selected diffusion and distance parameters. Highly connected or central nodes can distribute or accumulate signal across larger network regions than sparsely connected nodes. The simulation results show that the extent of propagated signal depends on the topological position of the experimentally measured nodes. Therefore, sensitivity analysis across α and $d_{max}$ values might be useful when applying the method to new datasets. Third, the simulations used in this work provided a controlled evaluation with known perturbed metabolites, but it does not capture the complexity, noise structure, and heterogeneity of real metabolomic datasets. Also, increasing the number of network derived features may affect downstream multivariate models. The use of identical CV partitions and permutation testing of $\Delta dQ^2$ can be used to assess the significance of changes in predictive performance linked to the propagation procedure. Finally, experimental validation, including targeted metabolite measurements and, where appropriate, complementary proteomic or transcriptomic measurements

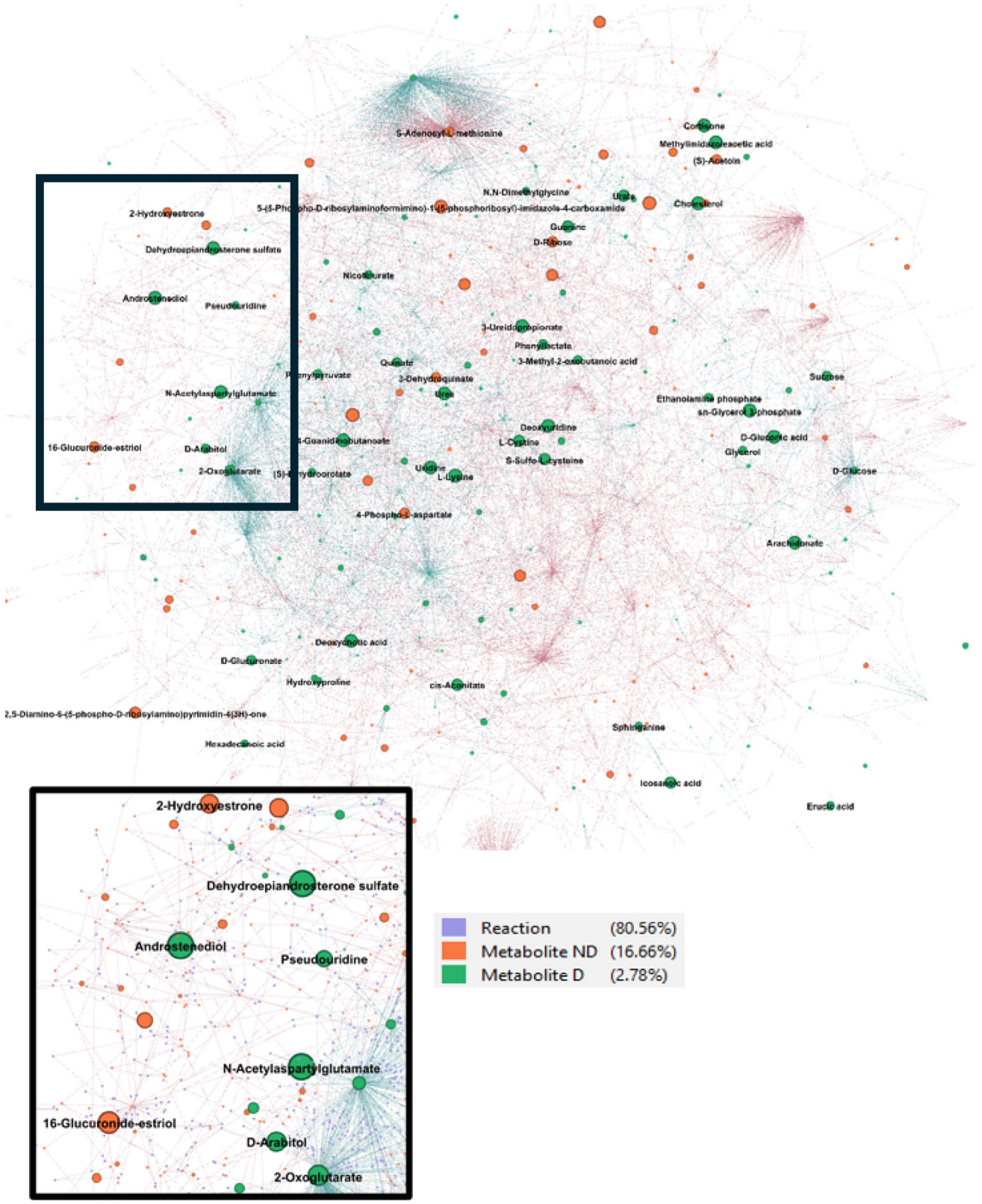


**Fig. 10** Section of the network showing the names of the 51 metabolites found significant (p-value < 0.05, SCR < 0.8 or > 1.25, and $max(\overline{SC_{E+P}}, \overline{SC_{E-only}}) > 0.05$) in the E-only vs. E + P analysis of the propagated bootstrap generated data set. Green nodes: initial non-zero metabolites, orange nodes: propagation-inferred metabolites, purple nodes: reactions. Metabolic node size is proportional to the −log p-value

of enzymes associated with prioritized reactions, might be required to validate the biological relevance of network-inferred candidates.

## 4 Conclusions

We present a network-based strategy for metabolomics analysis that integrates metabolite-level SCs with network topology through propagation. Under controlled simulation conditions, network propagation extended statistical signal to connected network regions, and improved discriminant performance while maintaining specificity in null comparisons. These results provide proof-of-concept that network propagation can support the exploration of partially observed metabolic systems by integrating local experimental data with a reference network structure. While the inferred SCs should be interpreted as hypothesis-generating and require independent experimental validation, this approach can be valuable in metabolomics, where incomplete coverage and annotation limit direct biological interpretation, and exploratory strategies are important in guiding downstream validation. Application to an experimental human metabolomics dataset demonstrated the feasibility of extending observed metabolomic signals to network-inferred candidate metabolites, while highlighting the exploratory nature of such inferred signals.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11306-026-02529-y.

**Acknowledgements** This work was supported by the Instituto de Salud Carlos III (ISC III), Spain, and co-funded by the European Union [grant number PI23/00202], and the Conselleria d'Educació, Cultura i Universitats, Generalitat Valenciana, Spain, under grant number CIAICO/2026/207. MMT acknowledges the grant RYC2021–031346-I, funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR, the project "PI24/00215" from Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union, and the Grant CNS2024–154299 funded by MICIU/AEI/10.13039/501100011033.

**Author contributions** JK: resources, analyzed data; MMT: resources, analyzed data; DPG: analyzed data; FAET: analyzed data; GQ: conceptualization, methodology, analyzed data. All authors wrote and reviewed the manuscript.

**Data availability** The data, the Gephi Network file, and the simulation results that support the findings of this study are available in Zenodo (https://zenodo.org/records/20808246 ). The experimental metabolomic data for this study was previously reported and described by Stevens et al. (2018) and it is deposited in the MetaboLights database with the identifier MTBLS136.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**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/.

## References

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment,* , Article P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Cabanski, C. R., Qi, Y., Yin, X., Bair, E., Hayward, M. C., Fan, C., Li, J., Wilkerson, M. D., Marron, J. S., Perou, C. M., & Hayes, D. N. (2010). SWISS MADE: Standardized within class sum of squares to evaluate methodologies and dataset elements. *PLoS One,5*, Article e9905. https://doi.org/10.1371/journal.pone.0009905

Cowen, L., Ideker, T., Raphael, B. J., & Sharan, R. (2017). Network propagation: A universal amplifier of genetic associations. *Nature Reviews Genetics*, *18*, 551–562. https://doi.org/10.1038/nrg.2017.38

Frainay, C., & Jourdan, F. (2017). Computational methods to identify metabolic sub-networks based on metabolomic profiles. *Briefings In Bioinformatics*, *18*, 43–56. https://doi.org/10.1093/bib/bbv115

Green, M. L., & Karp, P. D. (2006). The outcomes of pathway database computations depend on pathway ontology. *Nucleic Acids Research,34*, 3687–3697. https://doi.org/10.1093/nar/gkl438

Hui, T. X., Kasim, S., Md. Fudzee, M. F., Sutikno, T., Hassan, R., Aziz, I. A., Hasan, M. H., Jaafar, J., Alharbi, M., & Sen, S. C. (2023). A review of random walk-based method for the identification of disease genes and disease modules. *Ieee Access: Practical Innovations, Open Solutions,11*, 116366–116383. https://doi.org/10.1109/ACCESS.2023.3324985

Ivanovs, M., Kadikis, R., & Ozols, K. (2021). Perturbation-based methods for explaining deep neural networks: A survey. *Pattern Recognition Letters,150*, 228–234. https://doi.org/10.1016/j.patrec.2021.06.030

Kuligowski, J., Albiach-Delgado, A., Pérez-Guaita, D., & Quintás, G. (2025). Interpretation of multivariate metabolomic models through network-guided perturbation-based explanations. *Metabolomics*, *21*, 143. https://doi.org/10.1007/s11306-025-02347-8

Kuligowski, J., Moreno-Torres, M., & Quintás, G. (2024). Improving insights from metabolomic functional analysis combining multivariate tools. *Analytica Chimica Acta,1323*, Article 343062. https://doi.org/10.1016/j.aca.2024.343062

Lee, K. S., Su, X., & Huan, T. (2025). Metabolites are not genes — avoiding the misuse of pathway analysis in metabolomics. *Nature Metabolism*. https://doi.org/10.1038/s42255-025-01283-0

Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., Spigelman, A. F., MacDonald, P. E., Wishart, D. S., Li, S., & Xia, J. (2024). MetaboAnalyst 6.0: Towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkae253

Pirhaji, L., Milani, P., Leidl, M., Curran, T., Avila-Pacheco, J., Clish, C. B., White, F. M., Saghatelian, A., & Fraenkel, E. (2016). Revealing disease-associated pathways by network integration of untargeted metabolomics. *Nature Methods,13*, 770–776. https://doi.org/10.1038/nmeth.3940

Radrich, K., Tsuruoka, Y., Dobson, P., Gevorgyan, A., Swainston, N., Baart, G., & Schwartz, J. M. (2010). Integration of metabolic databases for the reconstruction of genome-scale metabolic networks. *Bmc Systems Biology*, *4*, 114. https://doi.org/10.1186/1752-0509-4-114

Sidak, D., Schwarzerová, J., Weckwerth, W., & Waldherr, S. (2022). Interpretable machine learning methods for predictions in systems biology from omics data. *Frontiers in Molecular Biosciences,9*, 926623. https://doi.org/10.3389/fmolb.2022.926623

Stevens, V. L., Wang, Y., Carter, B. D., Gaudet, M. M., & Gapstur, S. M. (2018). Serum metabolomic profiles associated with postmenopausal hormone use. *Metabolomics: Official Journal of the Metabolomic Society,14*, 97. https://doi.org/10.1007/s11306-018-1393-1

Visonà, G., Bouzigon, E., Demenais, F., & Schweikert, G. (2024). Network propagation for GWAS analysis: A practical guide to leveraging molecular networks for disease gene discovery. *Briefings in Bioinformatics,25*, Article bbae014. https://doi.org/10.1093/bib/bbae014

Westerhuis, J. A., van Velzen, E. J. J., Hoefsloot, H. C. J., & Smilde, A. K. (2008). Discriminant Q2 (DQ2) for improved discrimination in PLSDA models. *Metabolomics,4*, 293–296. https://doi.org/10.1007/s11306-008-0126-2

Wieder, C., Frainay, C., Poupin, N., Rodríguez-Mier, P., Vinson, F., Cooke, J., Lai, R. P., Bundy, J. G., Jourdan, F., & Ebbels, T. (2021). Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis. *PLoS Computational Biology,17*, Article e1009105. https://doi.org/10.1371/journal.pcbi.1009105

Zhou, D., Bousquet, O., Lal, T. N., Weston, J., & Schölkopf, B. (2003). Learning with local and global consistency. *Advances in Neural Information Processing Systems,16*, 321–328.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.