

Uncovering Protein Networks in Cardiovascular Proteomics

Authors

Maria Hasman, Manuel Mayr, and Konstantinos Theofilatos

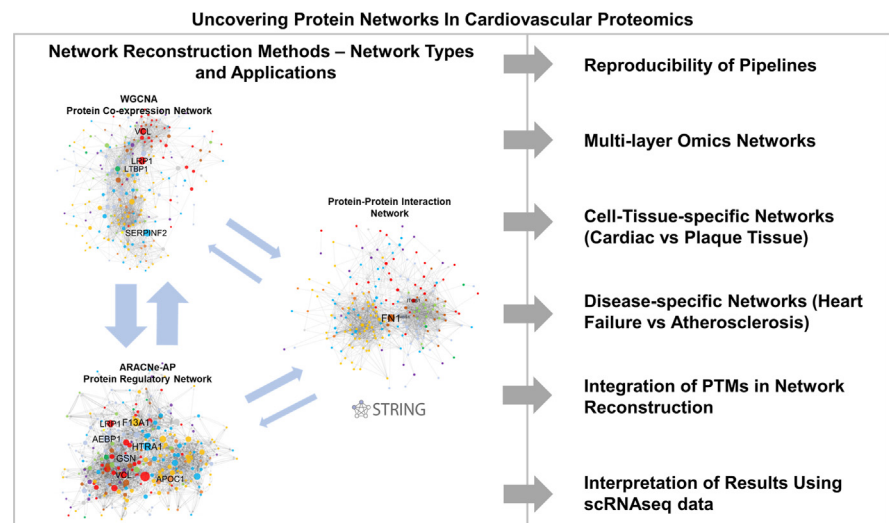
Correspondence

konstantinos.theofilatos@kcl.ac.uk

Graphical Abstract

In Brief

In the present article, we review and critically assess existing methods for reconstructing protein networks and discuss when each method is preferred for applications in cardiovascular research. We demonstrate the necessity to reconstruct networks separately for each cardiovascular tissue type and disease entity and provide illustrative examples of the importance of taking into consideration relevant post-translational modifications. Finally, we demonstrate and discuss how the findings of protein networks could be interpreted using single-cell RNA-sequencing data.



Highlights

- Co-expression and PPI networks are widely used in cardiovascular research.
- Directed regulatory networks are more suitable for generating causality hypotheses.
- Tissue, cell, and disease status triggers rewiring of protein networks.
- Considering post-translational modifications can refine the reconstructed networks.
- Integrating protein networks with other omics data enables scientific deductions.



Uncovering Protein Networks in Cardiovascular Proteomics

Maria Hasman, Manuel Mayr^{1b}, and Konstantinos Theofilatos^{1*}

Biological networks have been widely used in many different diseases to identify potential biomarkers and design drug targets. In the present review, we describe the main computational techniques for reconstructing and analyzing different types of protein networks and summarize the previous applications of such techniques in cardiovascular diseases. Existing tools are critically compared, discussing when each method is preferred such as the use of co-expression networks for functional annotation of protein clusters and the use of directed networks for inferring regulatory associations. Finally, we are presenting examples of reconstructing protein networks of different types (regulatory, co-expression, and protein-protein interaction networks). We demonstrate the necessity to reconstruct networks separately for each cardiovascular tissue type and disease entity and provide illustrative examples of the importance of taking into consideration relevant post-translational modifications. Finally, we demonstrate and discuss how the findings of protein networks could be interpreted using single-cell RNA-sequencing data.

Biological networks have been widely used in many different diseases to identify potential biomarkers, causal genes, and drug targets (1). Genetic interaction networks (2, 3) have been widely used to map genetic mutations with phenotypic changes, but the focus of the present article is on protein network reconstruction from quantitative data. The basic types of protein networks are the experimentally or *in silico* reconstructed protein-protein interaction (PPI) networks and the functional networks, which could show similar protein co-regulation, expression, or function. The latter can be split into regulatory and co-expression networks, according to whether their edges are directed or not. Networks can reveal useful biological and molecular information by inspecting two different and complementary types of network properties, the topological and the functional ones (4). Topological characteristics are used to represent the structural features of the network and are associated with biological properties and certain parameters, such as the betweenness centrality, which is used to reveal critical nodes. The functional approach clusters the nodes based on their functional information, such as cell

compartments and molecular functions (5). Recently, a systems-level approach, which involves focusing on a group of genes or proteins rather than on individual molecules, is being used to find mechanisms of complex diseases, which involve groups of genes or proteins (6). Several reviews are describing the different types of biological networks and analysis. Hu *et al.* (7) outlined different computational methods for identifying PPI networks. Vella *et al.* (8) described PPI and co-expression network reconstruction and analysis methods as well as studies involving the use of proteomics co-expression networks. Liu *et al.* (9) described different ways to identify critical nodes, and Meng *et al.* (10) described the different topological properties of PPI networks. Furthermore, previous reviews (11–15) have described the application of network analysis in diseases as well as systems biology approaches which include proteomics integration with other -omics technologies, network analysis, and their application in cardiovascular diseases (16, 17). To the best of our knowledge, however, none of them has focused on the different network types, the technical aspects of the network reconstruction methods, nor have they studied the effect of post-translational modifications and cell and tissue composition in the reconstructed networks. In this review, we discuss the role and types of protein networks, the different network analysis techniques, and tools, and focus on their application to tissue proteomics of clinical samples.

EXISTING METHODS FOR PROTEIN NETWORKS RECONSTRUCTION AND ANALYSIS

The basic steps of network analysis involve the reconstruction of PPI networks of undirected protein co-expression networks, and of directed protein regulatory networks, their clustering to identify significant modules, and their analysis and visualization to reveal key hub proteins that can serve as diagnostic, prognostic, or therapeutic biomarkers. PPI networks are static networks since they do not change, have the same connections between the nodes in all conditions, do not take into account the data of each experiment or different tissue or disease definition, and form a static integrated picture of protein activity. Table 1 presents the basic categories for protein network reconstruction and analysis, some

From the King's British Heart Foundation Centre, Kings College London, London, United Kingdom

*For correspondence: Konstantinos Theofilatos, konstantinos.theofilatos@kcl.ac.uk.

TABLE 1
Existing methods for the reconstruction and analysis of proteomic networks, indicative tools, and their implementations

Analysis type	Subcategory	Indicative methods	Advantages	Disadvantages	Implementation	Input data
Protein-Protein Interaction Network Reconstruction	Computationally and experimentally verified	STRING (21), Metascape (22)	<ul style="list-style-type: none"> i) Interactions for many species ii) Interactions from many sources iii) User-friendly network visualization and analysis 	<ul style="list-style-type: none"> High false positive rate Not full coverage of interactomes 	Webtool/ Standalone (R, Bioconductor Package)	Single or multiple gene list
Protein co-expression networks	Correlation-based	WGCNA (24)	<ul style="list-style-type: none"> i) Sparse Networks ii) Optimized Threshold 	<ul style="list-style-type: none"> i) One threshold for all nodes ii) Linear Associations Only iii) Undirected Networks 	Standalone (R/ Python)	<ul style="list-style-type: none"> i) Gene expression dataset ii) External trait data (eg clinical) to relate modules with
	Mutual Information Based	MIDER (25)	Distinguishes between direct and indirect links	Undirected Networks	Standalone (Matlab)	Time-series data
	Probabilistic	SEC (28)	Creates sparse matrices	The threshold needs to be decided with trial and error	Undirected Networks	Standalone (Matlab)
Protein Regulatory Networks	Mutual Information Based	ARACNe-AP (29)	<ul style="list-style-type: none"> i) Sparse Networks ii) Directed Networks iii) Non-linear associations iv) Removes indirect links 	<ul style="list-style-type: none"> i) Cannot discriminate between ii) positive and negative associations iii) Requires a given set of transcription factors 	Standalone (JAVA)	Gene expression data and tf/ gene list
	Probabilistic	BNW (27)	<ul style="list-style-type: none"> i) Handles noise and uncertainty ii) Directed Networks 	<ul style="list-style-type: none"> i) Does not support large networks ii) Feedback loops are not allowed iii) Static networks of ≤ 19 variables 	Webtool	<ul style="list-style-type: none"> i) Hybrid datasets contain both continuous (eg gene expression data) and discrete (eg genotypes) variables ii) Can include prior knowledge
	Machine-learning based	dynGENIE3 (33)	<ul style="list-style-type: none"> i) Directed Networks ii) Good scalability 	Semi-parametric	Standalone (Python/ Matlab/R)	<ul style="list-style-type: none"> i) Steady -state and time series expression data ii) Can include prior knowledge (eg known TFs)
Network Clustering	Hard-clustering	HipMCL (88)	<ul style="list-style-type: none"> i) Fast Clustering in Large Networks ii) Supports edge-weighted graphs clustering 	<ul style="list-style-type: none"> i) Does not allow unclustered nodes ii) Does not allow overlapping clusters 	Standalone (C++)	Not Applicable
	Soft-clustering	ClusterOne (39)	<ul style="list-style-type: none"> i) Allows overlapping clusters ii) Allows for unclustered nodes iii) Does not allow small clusters 	Lower Interpretability	Standalone (JAVA)/ Cytoscape plug-in	Not Applicable

STRING is included in the computational methods of PPI network reconstruction as it uses an in-silico probabilistic method for PPI predictions.

indicative tools, and their advantages and disadvantages. PPIs can be divided into physical (direct) interactions and indirect interactions and both types of networks can be either reconstructed with experimental methods or predicted using *in silico* machine learning, mathematical modeling, or other methods (4, 12). There are many databases, where experimentally verified PPI networks can be retrieved, such as iRefIndex (18), combining information from primary databases, such as IntAct (19) and BioGRID (20). PPIs can also be computationally predicted (STRING (21) and Metascape (22)). Such databases have been extensively used for the reconstruction of PPI networks in several diseases (10, 23).

The existing methods for co-expression network reconstruction were grouped into four basic categories: correlation-based, information theory-based, mathematical modeling, and other methods. Weighted gene correlation network analysis (WGCNA) (24) is one of the most widely used correlation-based approaches to construct co-expression networks. WGCNA provides tools to construct networks, identify modules, determine topological attributes, simulate, and visualize data. Correlation-based methods cannot model nonlinear associations. Thus, mutual information-based techniques, such as MIDER (25), have been introduced to address this issue. The problem of reconstructing co-expression networks is a mathematical optimization problem and this is the reason why mathematical and probabilistic modeling techniques, such as the Bayesian ones (26, 27) and the Sparse Estimation of the Correlation matrix (SEC) (28) have been widely used. However, these techniques suffer from a high number of assumptions and are not applicable to large-scale networks.

Protein regulatory networks are networks incorporating directionality and interaction type for each interaction. The existing methods for regulatory network reconstruction were grouped into three basic categories: information theory-based, mathematical modeling, and machine learning-based methods. One of the most widespread methods from this category is the Algorithm for the Reconstruction of Accurate Cellular Networks using adaptive partitioning strategy (ARACNe-AP) (29), which uses an information theoretic context based on the Data Processing Inequality (DPI) theorem to infer direct regulatory relations among transcriptional regulator proteins and target genes. ARACNe-AP estimates the mutual information threshold, uses bootstraps to reconstruct networks, and finally constructs a consensus network. Artificial Neural Networks (ANNs) can recognize any input pattern entered and create models of the data structure and connections that occurred during the procedure. Recurrent neural networks are widely presented as the most effective neural network-based models for gene regulatory network construction because of their ability to represent and model feedback and memory mechanisms (30). Rubiolo *et al.* (31) have used the Extreme Learning Machine (ELM)-supervised neural model to reconstruct regulatory networks from time series. Machine learning models have also been suggested to

reconstruct protein regulatory networks. Such an example is Gene Network Inference with Ensemble of trees (GENIE3) (32), an unsupervised method for network inference based on regression trees. This method is scalable, suitable for nonlinear data, and efficient in the case of many features. However, it only takes steady-state data as input. Thus, Huynh-Tu and Geurts (33) proposed an adaptation, the dynamic GENIE3 (dynGENIE3), which can combine both steady-state and time series data as an input. Keyl *et al.* (34) used a hybrid method that combines neural networks and a machine learning technique, layer-wise relevance propagation (LRP), to predict patient-level proteomic networks. Their method first uses neural networks to predict protein abundances and after training the network, it uses LRP to identify the input variables of the neural network that contributed the most to the prediction of the output.

Clustering is the first step after constructing and visualizing the network (with the most popular and widely used tool for visualization being Cytoscape (35)). Network clustering is used to group proteins with similar expression patterns and results into groups of nodes that often correspond to different functional groups. Clustering methods can be separated into three categories: (1) Hard clustering techniques, such as Markov Clustering (MCL) (36), and Restricted Neighborhood Search Clustering Algorithm (RNSC) (37), where all proteins are clustered and each protein can only belong to one cluster; (2) Soft clustering techniques, such as Molecular Complex Detection (MCODE) (38), ClusterONE (39) and WRNSC (40), where clusters can overlap and some proteins could remain unclustered; (3) Clustering techniques that are based on modularity density metric to detect communities, such as Louvain Clustering (41, 42). The Louvain clustering algorithm has found widespread use in various forms and has been applied to numerous biological networks. However, modularity-based algorithms like Louvain have a limitation in that they tend to exclude communities below a certain size threshold, known as the resolution limit. To overcome this limitation, a recent publication (43) used a recursive hierarchical approach in combination with the Louvain community detection algorithm to enable the resolution of hierarchically nested structures. Hierarchical clustering methods belong to the hard clustering techniques but are also able to provide a hierarchical organization among the proteins of a protein network (44). When all these network clustering methods are applied to physical PPI networks, the predicted clusters correspond to protein complexes (7). However, recently, these techniques for protein complex prediction were complemented with methods taking into consideration computationally predicted protein structure (45) or being applied directly to raw mass spectrometry data (46). Furthermore, another group of methods has been introduced to hierarchically organize networks (47) and to hierarchically organize tissues based on their underlying proteomic networks (48).

Another important analysis step involves the identification of essential nodes, also called hubs. Hub nodes are central nodes to the network and could be separated into intra- and intermodular hubs. Intra-modular hubs are central to each network cluster they belong to, whereas inter-modular hubs are central to the whole network (13). The most widely used methods for identifying hubs in a network are based on topological centralities such as the betweenness, degree, PageRank, and clustering coefficient with many tools, including Cytoscape (35), Network Analysis Tools (NeAT) (49) and NetConfer (50) enabling such analysis.

NETWORK APPROACHES FOR TISSUE AND DISEASE-SPECIFIC NETWORKS

The PPI databases can provide only a static form of the proteome interactome and not the cell- and tissue-specific context-dependent proteome connectivity (51). Cell-specific interaction networks can reveal characteristic biological properties and unique interactions of each cell, which are often related to distinct phenotypes (52). It has been observed among different cell-specific networks that, even when their interactomes share most of the proteins, the interactions among them differ. These interactome differences reveal cell-specific processes and are a result of distinct mechanisms, such as differential protein abundance, protein localization, or post-translational modifications (52–54). This cellular diversity leads to tissue-scale protein interactions. Attempts to report tissue-specific protein interactomes include TissueNet (55) a web tool that contains tissue-specific PPIs for human proteins and displays the expression of a query protein and all its interaction partners in the different tissues available. In the context of disease, Basha *et al.* (56) used MyProteinNet (57) to show tissue selectivity for gene expression and interactomes, exploring various tissues and including heart tissue. Greene *et al.* (53) used a kidney-specific network reconstructed with the NetWAS tool to train a classifier that can identify tissue-specific connectivity patterns associated with hypertension.

Figure 1 demonstrates a visual example of reconstructing the basic types of networks presented in Table 1 (PPI, protein co-expression, and protein regulatory networks). We have chosen a proteomics dataset for the matrisome of the left anterior descending coronary human arteries, using the most widely used method from each category. The matrisome was defined as the ensemble of Extracellular Matrix (ECM) and ECM-associated proteins. The ECM is a three-dimensional structure present in all tissues but different for each organ. It is composed of proteins such as proteoglycans, collagens, and glycoproteins (58). ECM remodeling in atherosclerosis plays an important role in plaque destabilization and progression (59). Bayesian methods and other mathematical modeling methods were not used since the examined tools (Table 1) were not supporting datasets of this sample and protein markers size. Limited overlap was found between the static PPI network and the reconstructed networks, as the

percentages of confirmed interactions of the PPI network against both the protein co-expression (8.47%) and the protein regulatory (9.11%) network was less than 10%. PPI networks, even the ones based on experimental evidence, are being created based on evidence from different types of tissues and conditions, and it is highly likely that most of them are not relevant to a particular tissue. As shown in Figure 1, significant hub proteins (betweenness centrality over 0.05) in the PPI network, such as Fibronectin 1 (FN1), were not returned as significant in the other networks. Protein co-expression and protein regulatory networks presented higher overlap, with more than 30% of the interactions of one network confirmed in the other (31.99% and 32.42% respectively). Moreover, significant overlap was observed in the hub proteins, with the membrane-associated proteins Vinculin (VCL) and LRP1 (LDL Receptor Related Protein 1) being hubs (betweenness centrality over 0.05) for both protein co-expression and regulatory networks. Further experimental validation should be performed to verify interactions and significant proteins experimentally by performing pull-down assay experiments with the revealed hub proteins and their candidate interaction partners. Moreover, instead of just exploring the static PPI networks, this analysis should be complemented with reconstruction and analysis of co-expression and regulatory networks in the specific tissues of interest to probe real interactions and disease mechanisms.

To illustrate the tissue and disease specificity of the reconstructed networks, we reconstructed two networks of the extracellular matrisome using label-free proteomics data from samples of previously published studies on ischemic heart failure (60) (Fig. 2) and atherosclerotic carotid plaques (61) (Fig. 3). For the reconstruction of these networks, we used only extracellular proteins that were consistently quantified in both datasets (<30% missing values per protein). Both networks were reconstructed using the ARACNe-AP information theory-based method with default parameters for network reconstruction (100 bootstraps for creating the network and Bonferroni correction of the nominal *p*-values of each inferred edge). With the network reconstruction, we were able to confirm known interactions in the heart tissue network such as the ones between collagens (COL6A1, COL6A2, COL6A3) and the ones between laminins (LAMA5, LAMB2, LAMC1). We performed network analysis for the two networks, using the NetworkAnalyzer (62) Cytoscape plug-in. The power law distribution was fitted to both networks' degree distribution, verifying the scale-free topology of the networks. The networks for plaque and heart tissue had comparable clustering coefficients of 0.24 and 0.27 respectively, suggesting a similar topology for both networks. However, the two disease- and tissue-specific networks differed in several other properties. For example, the heart tissue network has higher connectivity and a smaller number of sub-networks than the carotid plaque network, with 11 connected components and a radius of 5, compared to 23 connected components and a radius of 1.

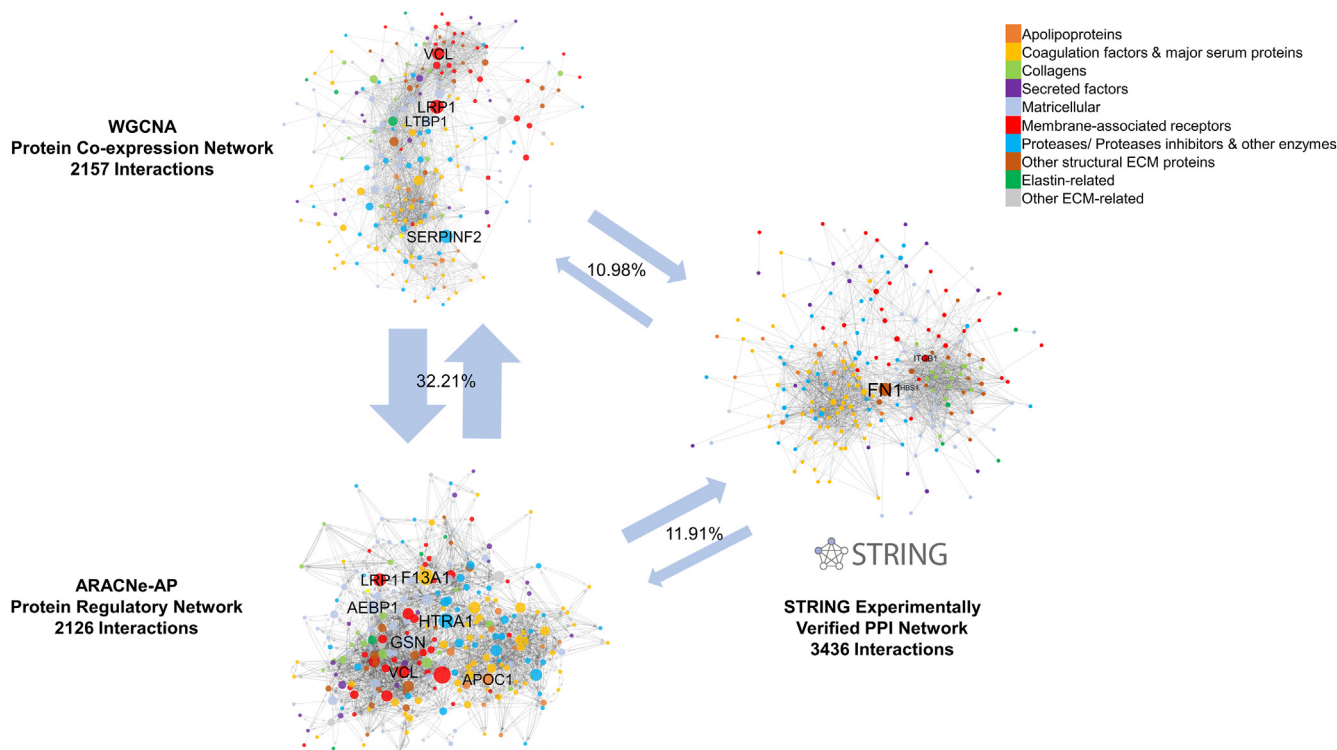


FIG. 1. Different networks to capture different aspects of the matrisome network of atherosclerotic plaques. Example of different types of networks for the matrisome of human *left* anterior descending coronary artery. Data-independent acquisition mass spectrometry data of 99 samples were used from the Parker *et al.* study (80) and filtered to keep only matrisome-related proteins, according to a custom matrisome database composed of extracellular proteins from the MatrisomeDB (81), apolipoproteins, and other secreted proteins that are consistently quantified (less than 20% missing values). The experimentally verified protein-protein interaction network was created by mining matrisome interactions from the STRING web tool (21). The protein co-expression network was reconstructed using the WGCNA pipeline (24), with Pearson's correlation as the interaction metric and 0.5 with a soft power of 10 as a threshold to infer interactions. ARACNe-AP (29) with default parameters was used to reconstruct the protein regulatory network with the same data. All networks were visualized using Cytoscape (35), proteins were colored based on the matrisome group they belong to, and node size was set to be proportional to the betweenness centrality of the node in the network. Only the nodes with betweenness centrality above 0.05 were labeled. The confirmed interactions of a network against another type of network are depicted with *arrows* connecting the different networks, and the average percentages (%) of common interactions between the two types of networks are also depicted. AEBP1, Adipocyte enhancer-binding protein one; APOC1, Apolipoprotein C-I; F13A1, Coagulation factor XIII A chain; FN1, Fibronectin; GSN, Gelsolin; HTRA1, Serine protease HTRA1; ITGB1, Integrin beta-1; LRP1, Prolow-density lipoprotein receptor-related protein one; LTBP1, Latent-transforming growth factor beta-binding protein one; SERPINF2, Alpha-2-antiplasmin; THBS1, Thrombospondin-1; VCL, Vinculin.

These differences suggest that the two networks have distinct structures and may be governed by different underlying biological processes, in particular, higher connectivity among the matricellular proteins in the ischemic heart tissue network compared to carotid plaques. In contrast, serum proteins and proteases are more central in the carotid plaque network since carotid endarterectomy lesions are expected to contain more serum-derived proteins than myocardial tissue. Clustering analysis (using the clusterONE Cytoscape plug-in, a minimum number of five proteins per cluster, and the MI) also verified this, with certain matricellular proteins forming a unique cluster for the cardiac tissue network that was not present in the carotid plaque network.

The most significant difference is observed in the top hub proteins using the degree centrality of each network.

Complement Factor H (*CHF*), Fibronectin (*FN1*), and Target of Nesh-SH3 (*ABI3BP*) are the top three interconnected proteins of the heart tissue network (with their degree centrality being 40, 23, and 22 respectively). Publicly available single-cell RNA-sequencing data from heart tissue samples (Fig. 4A, ExpressHeart web portal) showed that these matricellular and structural ECM proteins are mostly expressed in fibroblast and myofibroblast cells in the heart. It is noteworthy that cardiomyocytes were not included in this analysis, as the ExpressHeart web portal is only based on scRNAseq and does not include single nuclei data, thus cardiomyocytes are too big to detect with this technique. On the contrary, these proteins lost their central role in the carotid plaque network and stopped being hubs (Fig. 3), having a very smaller degree centrality, belonging to different subnetworks and are

Reconstructed CARDIAC extracellular protein network:

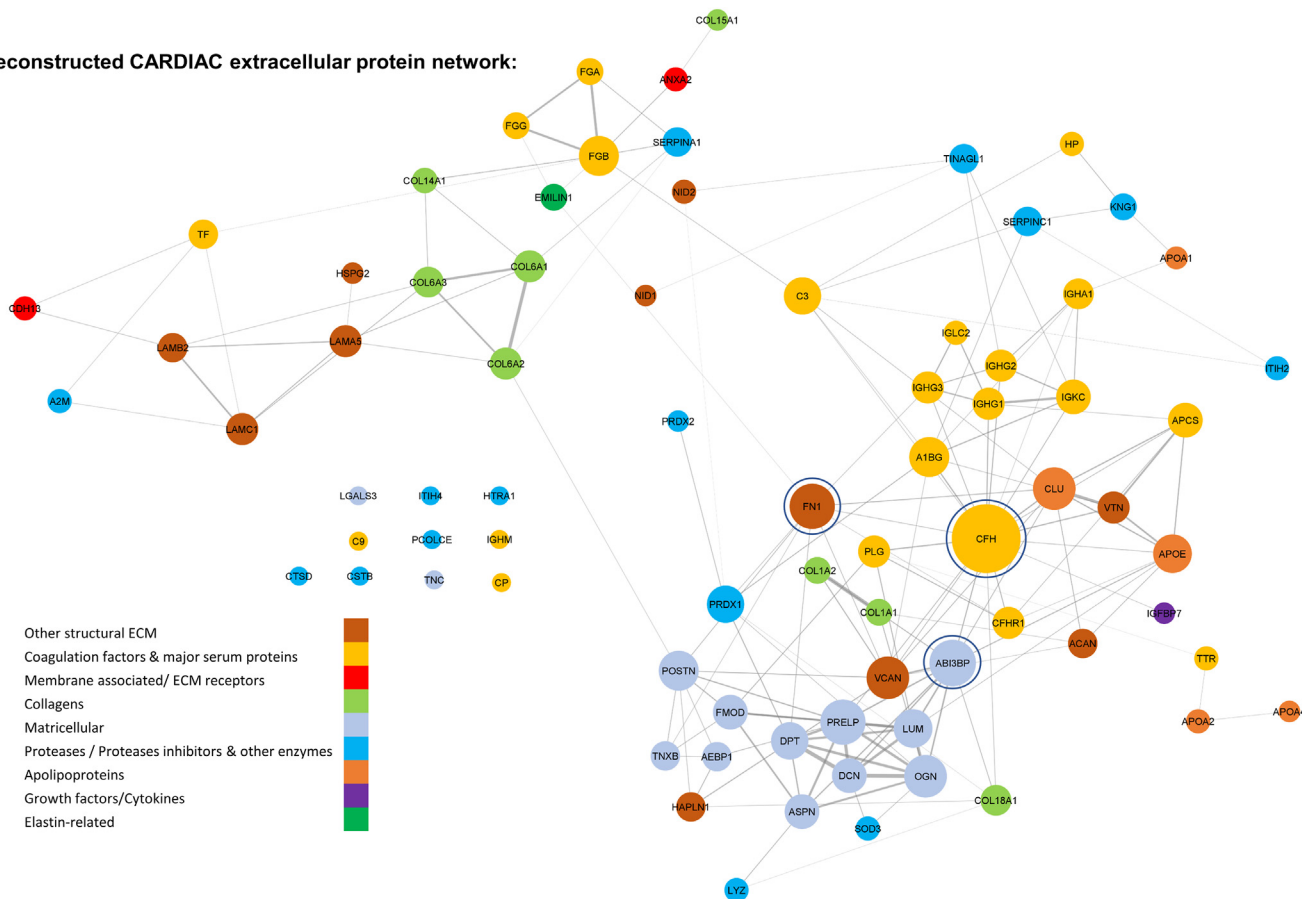


FIG. 2. **Heart tissue matrisome network.** Label-free discovery mass spectrometry data of 65 ischemic heart tissue samples were used from the study by Barallobre-Barreiro *et al.* (60) and filtered to keep only matrisome-related proteins, according to a custom matrisome database composed of extracellular proteins from the MatrisomeDB (81), apolipoproteins, and other secreted proteins that are consistently quantified (less than 30% missing values). ARACNe-AP (29) with default parameters was used to reconstruct the regulatory network, filtering out negative associations using the SIREN algorithm (82). Networks were visualized using Cytoscape (35), matrisome proteins were colored based on the functional category they belong to, edge width was set to be proportional to mutual information metric, node size was set to be proportional to its degree centrality and hub proteins are highlighted in blue.

expressed in both smooth muscle and endothelial cells (Fig. 4B). Immunoglobulins (Immunoglobulin Kappa Constant: *IGKC*, Immunoglobulin Heavy Constant Alpha 1: *IGHA1*) and serpin family A member 1 (*SERPINA1*) were the top three hub proteins in plaques (with a degree of 11 for immunoglobulins and 10 for *SERPINA1* respectively) and were highly expressed in plasma cells and monocytes/macrophages, respectively (Fig. 4B). Moreover, proteins with similar functionality were more closely connected in the heart tissue network than the carotid plaques one. One such example is the matricellular proteins and especially the proteoglycans (*DPT*, *PRELP*, *LUM*, *DCN*, *OGN*, *ASPN*), which form a highly connected network component in the heart tissue network (Fig. 2), whereas plaques showed less interconnectivity and belonged to different subnetworks. In opposite to their cellular expression in the heart tissue (myofibroblasts and fibroblasts), in plaques these proteins were not only expressed in fibroblasts but also in endothelial and smooth muscle cells (Fig. 4B), reflecting the

higher cell heterogeneity of carotid plaques. Thus, the combination of network analysis and single-cell RNA-sequencing data verified that the known cell composition differences between the two tissues and diseases are reflected in the reconstructed networks and identified different hub proteins for each matrisome network.

INTEGRATION OF PTMs IN NETWORK RECONSTRUCTION

Post-translational modifications (PTMs) are playing an important role in the structure and function of proteins and should be taken into consideration when reconstructing biological networks using proteomics data. Some tools already take PTMs in network reconstruction into consideration, such as the iPTMnet (63).

To assess the influence of PTMs, we reconstructed a protein regulatory network for atherosclerotic plaques based on the same dataset from Figure 3 but considering oxidized and unoxidized forms of each protein as separate nodes. To

Reconstructed PLAQUE ECM protein network

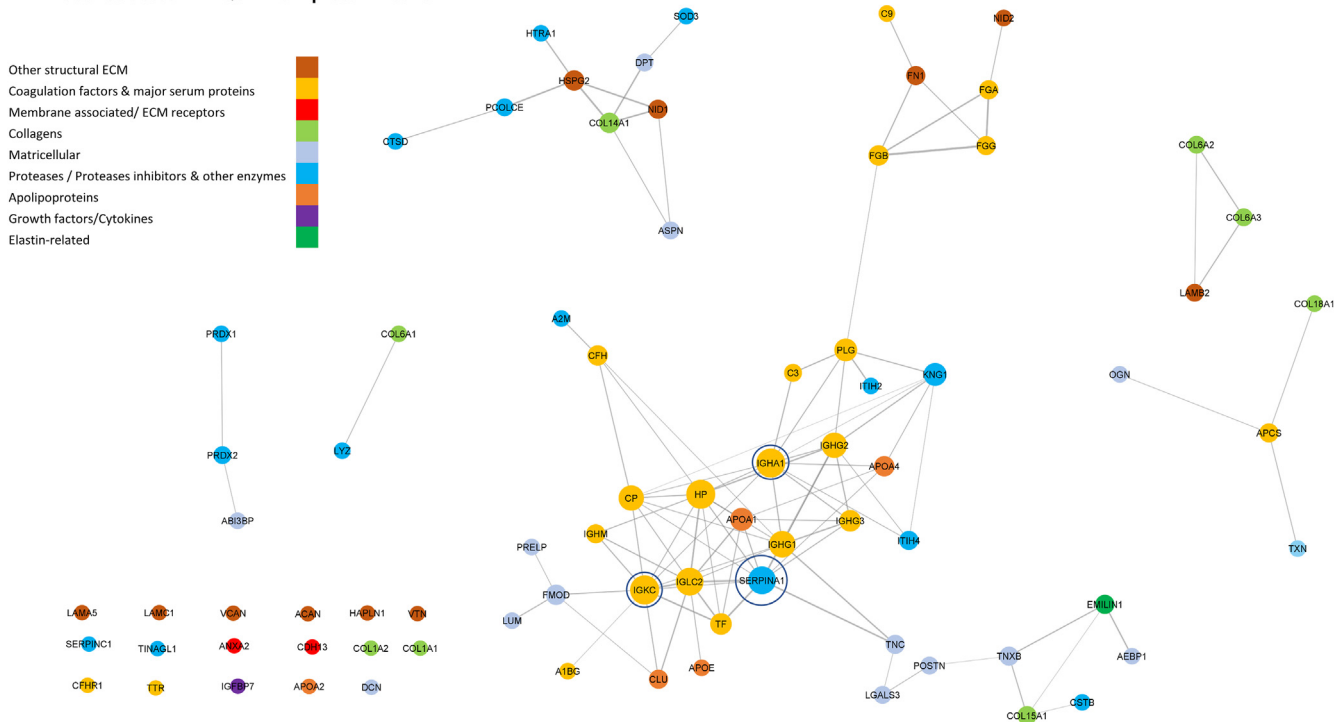


FIG. 3. **Carotid plaque matrisome network.** Label-free discovery mass spectrometry data of 12 carotid endarterectomy samples were used from the study by Langley *et al.* (61) and filtered to keep only matrisome-related proteins, according to a custom matrisome database composed of extracellular proteins from the MatrisomeDB (81), apolipoproteins, and other secreted proteins that are consistently quantified (less than 30% missing values). ARACNe-AP (29) with default parameters was used to reconstruct the regulatory network, filtering out negative associations using the SIREN algorithm (82). Networks were visualized using Cytoscape (35), matrisome proteins were colored based on the functional category they belong to, edge width was set to be proportional to mutual information metric and node size was set to be proportional to its degree centrality.

explore the effect of oxidation on the interactions and correlation patterns of the proteins in human atherosclerotic plaques, we exported the atherosclerotic dataset at the peptide level, choosing the oxidation of methionine, proline, and lysine residues as a dynamic modification. We used the consistently detected peptides (<30% missing values) to quantify the unmodified (at least two peptides per protein) and oxidized proteins. The WGCNA method (24) was used for network reconstruction. Filamin A (*FLNA*) was among the proteins which presented the biggest differentiation in its degree centrality when compared with its oxidized and unmodified form (54 and 34-degree centrality, respectively). Filamin A is a cellular protein that crosslinks actin filaments and links them to membrane glycoproteins and is involved in the remodeling of the cytoskeleton, to effect changes in cell shape and cell migration. Filamin A is also secreted, and it was consistently quantified in the extracellular extracts of the carotid plaque dataset (64). Seven *FLNA* peptides were found to be oxidized, with four of them being oxidized in methionine and three of them in proline amino acids. Accurate quantification of methionine oxidation in a proteomics scale has technical limitations as methionine could be readily oxidized during

sample preparation (65). However, the sum of the abundances of methionine oxidized peptides was highly correlated with the sum of proline oxidized peptides of *FLNA* (Spearman's Rho: 0.97). Thus, we included methionine peptides. As shown in Figure 4C, the oxidized and non-oxidized forms of *FLNA* share common interactors, but each form has also unique interactors. The unique interactors of oxidized *FLNA* are enriched in cellular responses related to cell-cell communication, neutrophil degranulation, and smooth muscle cell contraction.

MULTILAYER-OMICS NETWORKS

Exploring the interconnectivity of the different molecules that constitute cells, that is, proteins, genes, and metabolites, could enable capturing more complex cellular mechanisms in comparison to separately studying each omics modality. Hammoud and Kramer (66) have recently reviewed such networks and their application to PPI, cell, and gene expression networks. Such an example of a multi-omics network reconstruction software is COSMOS (67), which combines transcriptomics, phosphoproteomics, and metabolomics datasets to build a multi-omics network. Using prior knowledge, this

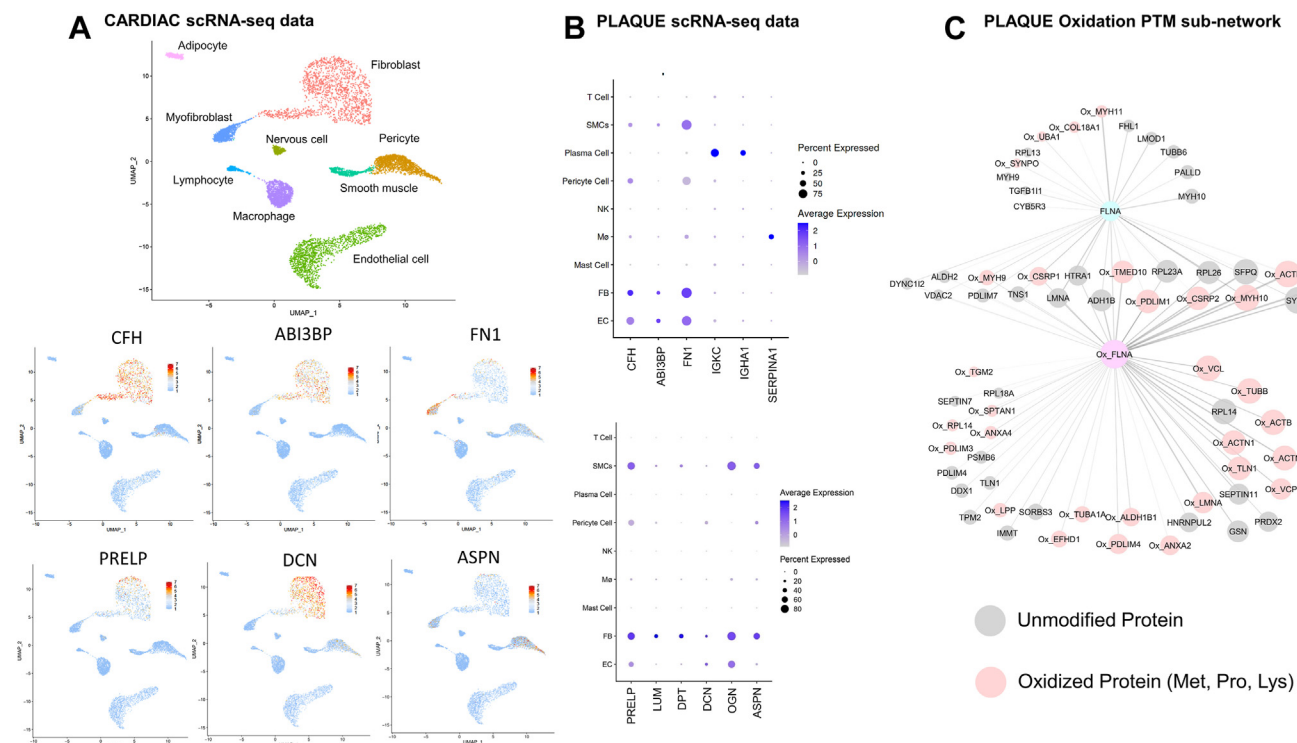


FIG. 4. Interpreting cardiac and plaque protein networks using scRNAseq data and the effect of oxidation PTM in the interactions of Filamin-A (FLNA). A, uniform manifold approximation (UMAP) feature plots of the expression of the *top* central nodes (hubs) of the network and other central proteoglycans, using the Hocker *et al.* single-cell RNA-seq dataset (83) of human heart tissue samples (8993 cells from two healthy donors) from the ExpressHeart web portal (84). B, expression dot plots of *top* three central proteins (hubs) in the carotid plaques network and the cardiac tissue network and other central in the cardiac protein network proteoglycans (Fig. 2) using the PlaQView web tool (85), scRNA-seq data from carotid plaques (Pan *et al.* dataset (86), n = 3) and Aran *et al.* method (87) to label cell clusters. C, label-free discovery mass spectrometry data of 12 carotid endarterectomy samples were used from the study of Langley *et al.* (61). Data was exported in peptide level from Proteome Discoverer software and oxidation of proline, lysine, and methionine was used as a dynamic modification. The network was reconstructed using the WGCNA method (24) with the Pearson’s correlation method, a first soft power of 10 and a final hard threshold of 0.25. Nodes are proportional to the degree of each protein (based on all proteins network), edge width is proportional to the MI metric, and node colors are set according to the oxidation modification (light grey for unmodified proteins and pink for oxidized proteins, with the unmodified and oxidized form of FLNA having light blue and light purple color respectively). EC, Endothelial cells; FB, Fibroblast cells; Mo, Monocytes/macrophages; NK, Natural Killer cells; SMCs, Smooth Muscle Cells.

method first builds networks from omics datasets from public databases, then refines those networks by removing the edges that create incorrect predictions when applied to only one omics modality and, finally, further filters the network based on only differentially expressed genes, proteins, and metabolites. The resulting network consists of the differentially expressed molecules and their interactors in a set number of regulatory steps away, creating a causal network.

REPRODUCIBILITY AND METHODS SHARING

As highlighted in Table 1, various methods are available for reconstructing different types of networks using proteomics data. These methods often require different programming languages, input and output file formats, and software package versions. In addition, different operating system requirements can further complicate the use of these tools and the reproducibility of results and analysis. To address this

problem and ensure reproducibility, Netbooks (68) have been developed either based on Jupyter Notebooks (69) or other markdown formats, including R Markdown (70). Jupyter Notebook is a web-based, open-source application originally designed for Python programming, but it now supports over 50 programming languages. With Jupyter Notebook, users can create and share documents that contain text, code, and other project-related materials. These documents, known as notebooks, can include various types of output, such as plots, interactive graphics, and other visualizations, and can be easily manipulated using third-party software tools.

CONCLUSION

The reconstruction and study of different types of disease networks, as demonstrated in this review article, is mostly based on the reconstruction of co-expression networks using established pipelines such as WGCNA, or on the use of static

commercial (e.g. IPA) or publicly available (e.g. STRING) PPI networks. Despite their simplicity and ease of use, they have their limitations. Many researchers try to overcome this issue by combining proteomics with genomics data in an attempt to identify pQTL (71, 72), but these approaches are limited to genomic variants which alter protein abundances and, thus, fail to capture pathogenic mechanisms related to post-translational modifications, protein degradation, and protein-protein interactions. To overcome these limitations, more robust pipelines and methods are needed to reconstruct and study directed regulatory networks, which are more suitable for generating and validating causality hypotheses and performing simulations. Moreover, multi-omics networks combining RNA and protein networks (17, 73) and the integration of single-cell RNA-sequencing and proteomics data (74), further allow the parallel analysis of transcriptional and translational mechanisms, but this analysis requires the development of new methods, such as the consensus clustering (75) that has been recently introduced.

Another important aspect is the validation of these network findings. Network analysis is a strong tool for hypothesis generation but its application on direct or indirect interactions between proteins should be further validated with additional experimental techniques. Two-hybrid screening is a widely used method to detect binary interactions in eukaryotic cells. One of the main limitations of this technique is that it can generate false positives or false negatives. Other techniques for experimentally validating PPI networks at scale include MS-based techniques. In affinity purification (AP-MS) a bait protein is purified with antibodies along with its potential interactor partners and these purified proteins are then identified by MS. In cross-linking MS (XL-MS), a protein mixture is incubated with chemical cross-linkers and then analyzed by LC-MS/MS for cross-linked peptides. Further support can be derived from co-fractionation experiments coupled to MS, where samples are lysed, and fractionated, and each fraction is analyzed with LC-MS/MS and protein abundances are plotted across fractions. Co-fractionation, however, is no proof for genuine protein interactions but can be considered as supporting evidence for an interactome network (76). The development of cross-linking mass spectrometry technologies (77) has allowed for the validation of protein-protein interactions on large scale, while single-cell proteomics (78, 79) can provide an additional layer of validation for the findings inferred by combining protein networks and single-cell RNA-sequencing data.

The network reconstruction process is highly affected by covariates and medications should thus be taken into consideration when reconstructing networks and when interpreting findings based on them. Moreover, as protein interaction networks are significantly different among different diseases and sample types, with cell composition being one of the most important factors. Thus, it is paramount that networks should be reconstructed specifically for each

tissue type and disease entity, while when proteomics data are used for reconstructing the networks post-translational modifications should be taken into consideration as illustrated in the example provided in the present manuscript for carotid plaque networks. With all these factors taken into consideration, the comparisons of these networks become more meaningful with regard to identifying potential pathophysiological mechanisms that can then be further validated experimentally.

Funding and additional information—M. H. was supported by a BHF Centre Interdisciplinary non-clinical Ph.D. Studentship. M. M. is a British Heart Foundation (BHF) Chair Holder (CH/16/3/32406) with BHF programme grant support (RG/16/14/32397). K. T. and M. M. were also supported with a BHF project grant (PG/20/10387). M. M. is also supported by the Leducq Foundation (18CVD02) and the VASCage – Research Center on Vascular Ageing and Stroke (No. 898252).

Author contributions—M. H. formal analysis; M. H. writing—original draft; M. H and K. T. visualization; M. M and K. T. conceptualization; M. M. and K. T. Writing—Review & Editing; M. M. resources; M. M. and K. T. supervision; M. M. and K. T. funding acquisition; K. T. project administration.

Conflict of interest—The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Abbreviations—The abbreviations used are: ANNs, Artificial Neural Networks; ARACNe-AP, Algorithm for the Reconstruction of Accurate Cellular Networks using adaptive partitioning strategy; CHF, Complement Factor H; DPI, Data Processing Inequality; dynGENIE3, dynamic GENIE3; ECM, Extracellular Matrix; ELM, Extreme Learning Machine; FLNA, filamin A; FN1, Fibronectin 1; LRP, layer-wise relevance propagation; LRP1, LDL Receptor Related Protein 1; MCL, Markov Clustering; MCODE, Molecular Complex Detection; NeAT, Network Analysis Tools; PPI, protein-protein interaction; PTM, post-translational modification; RNSC, Restricted Neighborhood Search Clustering Algorithm; SEC, Sparse Estimation of the Correlation matrix; SERPINA1, serpin family A member 1; VCL, Vinculin; WGCNA, weighted gene correlation network analysis.

Received November 14, 2022, and in revised form, May 1, 2023
Published, MCPRO Papers in Press, June 24, 2023, <https://doi.org/10.1016/j.mcpro.2023.100607>

REFERENCES

1. Paci, P., Fisco, G., Conte, F., Wang, R. S., Farina, L., and Loscalzo, J. (2021) Gene co-expression in the interactome: moving from correlation toward causation via an integrated approach to disease module discovery. *NPJ Sys Biol. Appl.* 7, 1–11
2. Costanzo, M., VanderSluis, B., Koch, E. N., Baryshnikova, A., Pons, C., Tan, G., et al. (2016) A global genetic interaction network maps a wiring diagram of cellular function. *Science* 353, aaf1420

3. Costanzo, M., Kuzmin, E., van Leeuwen, J., Mair, B., Moffat, J., Boone, C., et al. (2019) Global genetic networks and the genotype-to-phenotype relationship. *Cell* **177**, 85–100
4. Miriyala, S. K., Anbarasu, A., and Ramaiah, S. (2018) Discerning molecular interactions: a comprehensive review on biomolecular interaction databases and network analysis tools. *Gene* **642**, 84–94
5. Iacobucci, I., Monaco, V., Cozzolino, F., and Monti, M. (2021) From classical to new generation approaches: an excursus of -omics methods for investigation of protein-protein interaction networks. *J. Proteomics* **230**, 103990
6. Schmidt, H. H. H. W., and Menche, J. (2022) The regulatory network architecture of cardiometabolic diseases. *Nat. Genet.* **54**, 2–3
7. Hu, L., Wang, X., Huang, Y. A., Hu, P., and You, Z. H. (2021) A survey on computational models for predicting protein-protein interactions. *Brief. Bioinform.* **22**, bbab036
8. Vella, D., Zoppis, I., Mauri, G., Mauri, P., and Di Silvestre, D. (2017) From protein-protein interactions to protein co-expression networks: a new perspective to evaluate large-scale proteomic data. *EURASIP J. Bioinform. Syst. Biol.* **2017**, 1–16
9. Liu, X., Hong, Z., Liu, J., Lin, Y., Rodríguez-Patón, A., Zou, Q., et al. (2020) Computational methods for identifying the critical nodes in biological networks. *Brief. Bioinform.* **21**, 486–497
10. Meng, X., Li, W., Peng, X., Li, Y., and Li, M. (2021) Protein interaction networks: centrality, modularity, dynamics, and applications. *Front. Comput. Sci.* **15**, 1–17
11. Sonawane, A. R., Aikawa, E., and Aikawa, M. (2022) Connections for matters of the heart: network medicine in cardiovascular diseases. *Front. Cardiovasc. Med.* **9**, 1174
12. Cho, D. Y., Kim, Y. A., and Przytycka, T. M. (2012) Chapter 5: network biology approach to complex diseases. *PLOS Comput. Biol.* **8**, e1002820
13. van Dam, S., Vösa, U., van der Graaf, A., Franke, L., and de Magalhães, J. P. (2018) Gene co-expression analysis for functional classification and gene-disease predictions. *Brief. Bioinform.* **19**, 575–592
14. Saint-Antoine, M. M., and Singh, A. (2020) Network inference in systems biology: recent developments, challenges, and applications. *Curr. Opin. Biotechnol.* **63**, 89–98
15. Zhang, P., and Iltan, Y. (2019) Biological network approaches and applications in rare disease studies. *Genes (Basel)* **10**, 797
16. Langley, S. R., Dwyer, J., Drozdov, I., Yin, X., and Mayr, M. (2013) Proteomics: from single molecules to biological pathways. *Cardiovasc. Res.* **97**, 612–622
17. Joshi, A., Rienks, M., Theofilatos, K., and Mayr, M. (2020) Systems biology in cardiovascular disease: a multiomics approach. *Nat. Rev. Cardiol.* **18**, 313–330
18. Razick, S., Magklaras, G., and Donaldson, I. M. (2008) iRefIndex: a consolidated protein interaction database with provenance. *BMC Bioinformatics* **9**, 1–19
19. Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., et al. (2014) The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–D363
20. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B. J., Stark, C., Willems, A., et al. (2021) The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* **30**, 187–200
21. Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021) The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurements sets. *Nucleic Acids Res.* **49**, D605–D612
22. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1–10
23. Bostanci, N., Selevsek, N., Wolski, W., Grossmann, J., Bao, K., Wahlander, A., et al. (2018) Targeted proteomics guided by label-free quantitative proteome analysis in saliva reveal transition signatures from health to periodontal disease. *Mol. Cell. Proteomics* **17**, 1392–1409
24. Langfelder, P., and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 1–13
25. Villaverde, A. F., Ross, J., Morán, F., and Banga, J. R. (2014) MIDER: network inference with mutual information distance and entropy reduction. *PLOS One* **9**, e96732
26. Noor, A., Serpedin, E., Nounou, M., Nounou, H., Mohamed, N., and Chouchane, L. (2013) An Overview of the Statistical Methods Used for Inferring Gene Regulatory Networks and Protein-Protein Interaction Networks. *Adv. Bioinform* **2013**, 953814
27. Ziebarth, J. D., Bhattacharya, A., and Cui, Y. (2013) Bayesian network webserver: a comprehensive tool for biological network modeling. *Bioinformatics* **29**, 2801–2803
28. Cui, Y., Leng, C., and Sun, D. (2016) Sparse estimation of high-dimensional correlation matrices. *Comput. Stat. Data Anal.* **93**, 390–403
29. Lachmann, A., Giorgi, F. M., Lopez, G., and Califano, A. (2016) ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. *Bioinformatics* **32**, 2233–2235
30. Chai, L. E., Loh, S. K., Low, S. T., Mohamad, M. S., Deris, S., and Zakaria, Z. (2014) A review on the computational approaches for gene regulatory network construction. *Comput. Biol. Med.* **48**, 55–65
31. Rubiolo, M., Milone, D. H., and Stegmayer, G. (2018) Extreme learning machines for reverse engineering of gene regulatory networks from expression time series. *Bioinformatics* **34**, 1253–1260
32. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010) Inferring regulatory networks from expression data using tree-based methods. *PLOS One* **5**, e12776
33. Huynh-Thu, V. A., and Geurts, P. (2018) dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. *Sci. Rep.* **8**, 1–12
34. Keyl, P., Bockmayr, M., Heim, D., Dernbach, G., Montavon, G., Müller, K. R., et al. (2022) Patient-level proteomic network prediction by explainable artificial intelligence. *NPJ Precis. Oncol.* **6**, 35
35. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504
36. Enright, A. J., Van Dongen, S., and Ouzounis, C. A. (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584
37. King, A. D., Pržulj, N., and Jurisica, I. (2004) Protein complex prediction via cost-based clustering. *Bioinformatics* **20**, 3013–3020
38. Bader, G. D., and Hogue, C. W. V. (2003) An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**, 1–27
39. Nepusz, T., Yu, H., and Paccanaro, A. (2012) Detecting overlapping protein complexes in protein-protein interaction networks. *Nat. Methods* **9**, 471–472
40. Dimitrakopoulos, C., Theofilatos, K., Pegkas, A., Likothanassis, S., and Mavroudi, S. (2016) Predicting overlapping protein complexes from weighted protein interaction graphs by gradually expanding dense neighborhoods. *Artif. Intell. Med.* **71**, 62–69
41. Blondel, V. D., Guillaume, J. L., Lambiotte, R., and Lefebvre, E. (2008) Fast unfolding of communities in large networks. *J. Stat. Mech. Theor. Exp* **2008**, P10008
42. Lu, H., Halappanavar, M., and Kalyanaraman, A. (2015) Parallel heuristics for scalable community detection. *Parallel Comput.* **47**, 19–37
43. Akiki, J., and Abdallah, G. (2019) Determining the hierarchical architecture of the human brain using subject-level clustering of functional networks. *Sci. Rep.* **9**, 1–15
44. Manipur, I., Giordano, M., Piccirillo, M., Parashuraman, S., and Maddalena, L. (2023) Community detection in protein-protein interaction networks and applications. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **20**, 217–237
45. Drake, Z. C., Seffernick, J. T., and Lindert, S. (2022) Protein complex prediction using Rosetta, AlphaFold, and mass spectrometry covalent labeling. *Nat. Commun.* **13**, 7846
46. Fossati, A., Li, C., Uliana, F., Wendt, F., Frommelt, F., Sykacek, P., et al. (2021) PCprophet: a framework for protein complex prediction and differential analysis using proteomic data. *Nat. Methods* **18**, 520–527
47. Kovács, I. A., Palotai, R., Szalay, M. S., and Csermely, P. (2010) Community landscapes: an integrative approach to determine overlapping network module hierarchy, identify key nodes and predict network dynamics. *PLOS One* **5**, e12528
48. Zitnik, M., and Leskovec, J. (2017) Predicting multicellular function through multi-layer tissue networks. *Bioinformatics* **33**, i190–i198
49. Brohée, S., Faust, K., Lima-Mendez, G., Sand, O., Janky, R., Vanderstocken, G., et al. (2008) NeAT: a toolbox for the analysis of biological

- networks, clusters, classes and pathways. *Nucleic Acids Res.* **36**, W444–W451
50. Nagpal, S., Baksi, K. D., Kuntal, B. K., and Mande, S. S. (2020) NetConfer: a web application for comparative analysis of multiple biological networks. *BMC Biol.* **18**, 1–12
 51. Bludau, I., and Aebersold, R. (2020) Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nat. Rev. Mol. Cell Biol.* **21**, 327–340
 52. Huttlin, E. L., Bruckner, R. J., Navarrete-Perea, J., Cannon, J. R., Baltier, K., Gebreab, F., et al. (2021) Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* **184**, 3022–3040
 53. Greene, C. S., Krishnan, A., Wong, A. K., Ricciotti, E., Zelaya, R. A., Himmelstein, D. S., et al. (2015) Understanding multicellular function and disease with human tissue-specific networks. *Nat. Genet.* **47**, 569–576
 54. Skinnider, M. A., Scott, N. E., Prudova, A., Kerr, C. H., Stoynov, N., Stacey, R. G., et al. (2021) An atlas of protein-protein interactions across mouse tissues. *Cell* **184**, 4073–4089.e17
 55. Basha, O., Barshir, R., Sharon, M., Lerman, E., Kirson, B. F., Hekselman, I., et al. (2017) The TissueNet v.2 database: a quantitative view of protein-protein interactions across human tissues. *Nucleic Acids Res.* **45**, D427–D431
 56. Basha, O., Argov, C. M., Artzy, R., Zoabi, Y., Hekselman, I., Alfandari, L., et al. (2020) Differential network analysis of multiple human tissue interactomes highlights tissue-selective processes and genetic disorder genes. *Bioinformatics* **36**, 2821–2828
 57. Basha, O., Flom, D., Barshir, R., Smoly, I., Tirman, S., and Yeger-Lotem, E. (2015) MyProteinNet: build up-to-date protein interaction networks for organisms, tissues and user-defined contexts. *Nucleic Acids Res.* **43**, W258–W263
 58. Raghunathan, R., Sethi, M. K., Klein, J. A., and Zaia, J. (2019) Proteomics, Glycomics, and Glycoproteomics of Matrisome Molecules. *Mol. Cell. Proteomics* **18**, 2138–2148
 59. Bonnans, C., Chou, J., and Werb, Z. (2014) Remodelling the extracellular matrix in development and disease. *Nat. Rev. Mol. Cell Biol.* **15**, 786–801
 60. Barallobre-Barreiro, J., Radovits, T., Fava, M., Mayr, U., Lin, W. Y., Ermolaeva, E., et al. (2021) Extracellular matrix in heart failure: role of ADAMTS5 in proteoglycan remodeling. *Circulation* **144**, 2021–2034
 61. Langley, S. R., Willeit, K., Didangelos, A., Matic, L. P., Skroblin, P., Barallobre-Barreiro, J., et al. (2017) Extracellular matrix proteomics identifies molecular signature of symptomatic carotid plaques. *J. Clin. Invest.* **127**, 1546–1560
 62. Doncheva, N. T., Assenov, Y., Domingues, F. S., and Albrecht, M. (2012) Topological analysis and interactive visualization of biological networks and protein structures. *Nat. Protoc.* **7**, 670–685
 63. Huang, H., Arighi, C. N., Ross, K. E., Ren, J., Li, G., Chen, S. C., et al. (2018) iPTMnet: an integrated resource for protein post-translational modification network discovery. *Nucleic Acids Res.* **46**, D542–D550
 64. Nakamura, F., Kumeta, K., Hida, T., Isono, T., Nakayama, Y., Kuramata-Matsuoka, E., et al. (2014) Amino- and carboxyl-terminal domains of Filamin-A interact with CRMP1 to mediate Sema3A signalling. *Nat. Commun.* **5**, 5325
 65. Bettinger, J. Q., Welle, K. A., Hryhorenko, J. R., and Ghaemmaghami, S. (2019) Quantitative Analysis of *in Vivo* methionine oxidation of the human proteome. *J. Proteome Res.* **19**, 624–633
 66. Hammoud, Z., and Kramer, F. (2020) Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Anal.* **5**, 1–18
 67. Dugourd, A., Kuppe, C., Sciacovelli, M., Gjerga, E., Gabor, A., Emdal, K. B., et al. (2021) Causal integration of multi-omics data with prior knowledge to generate mechanistic hypotheses. *Mol. Syst. Biol.* **17**, e9730
 68. Ben Guebila, M., Weighill, D., Lopes-Ramos, C. M., Burkholz, R., Pop, R. T., Palepu, K., et al. (2022) An online notebook resource for reproducible inference, analysis and publication of gene regulatory networks. *Nat. Methods* **19**, 511–513
 69. Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., et al. (2016). *Jupyter Notebooks-A Publishing Format for Reproducible Computational Workflows* (Vol. 2016,, 87–90
 70. Yihui Xie, J. J., and Allaire, G. G. (2018) *R Markdown: The Definitive Guide*. CRC Press, Boca Raton, FL
 71. Yao, C., Chen, G., Song, C., Keefe, J., Mendelson, M., Huan, T., et al. (2018) Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nat. Commun.* **9**, 1–11
 72. Zhang, J., Dutta, D., Köttgen, A., Tin, A., Schlosser, P., Grams, M. E., et al. (2022) Plasma proteome analyses in individuals of European and African ancestry identify cis-pQTLs and models for proteome-wide association studies. *Nat. Genet.* **54**, 593–602
 73. Jin, H., Goossens, P., Juhasz, P., Eijgelaar, W., Manca, M., Karel, J. M. H., et al. (2021) Integrative multiomics analysis of human atherosclerosis reveals a serum response factor-driven network associated with intra-plaque hemorrhage. *Clin. Transl. Med.* **11**, e458
 74. Marx, V. (2019) A dream of single-cell proteomics. *Nat. Methods* **16**, 809–812
 75. Jeub, L. G. S., Sporns, O., and Fortunato, S. (2018) Multiresolution consensus clustering in networks. *Sci. Rep.* **8**, 1–16
 76. Low, T. Y., Syafruddin, S. E., Mohtar, M. A., Vellaichamy, A., A Rahman, N. S., Pung, Y. F., et al. (2021) Recent progress in mass spectrometry-based strategies for elucidating protein-protein interactions. *Cell. Mol. Life Sci.* **78**, 5325–5339
 77. Barysz, H. M., and Malmström, J. (2018) Development of large-scale cross-linking mass spectrometry. *Mol. Cell. Proteomics* **17**, 1055–1066
 78. Brunner, A.-D., Thielert, M., Vasilopoulou, C., Ammar, C., Coscia, F., Mund, A., et al. (2022) Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol. Syst. Biol.* **18**, e10798
 79. Schoof, E. M., Furtwängler, B., Üresin, N., Rapin, N., Savickas, S., Gentil, C., et al. (2021) Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **12**, 1–15
 80. Parker, S. J., Chen, L., Spivia, W., Saylor, G., Mao, C., Venkatraman, V., et al. (2020) Identification of putative early atherosclerosis biomarkers by unsupervised deconvolution of heterogeneous vascular proteomes. *J. Proteome Res.* **19**, 2794–2806
 81. Shao, X., Taha, I. N., Clauser, K. R., Gao, Y. T., and Naba, A. (2020) MatrisomeDB: the ECM-protein knowledge database. *Nucleic Acids Res.* **48**, D1136–D1144
 82. Khosravi, P., Gazestani, V. H., Pirhaji, L., Law, B., Sadeghi, M., Goliaei, B., et al. (2015) Inferring interaction type in gene regulatory networks using co-expression data. *Algorithms Mol. Biol.* **10**, 1–11
 83. Hocker, J. D., Poirion, O. B., Zhu, F., Buchanan, J., Zhang, K., Chiou, J., et al. (2021) Cardiac cell type-specific gene regulatory programs and disease risk association. *Sci. Adv.* **7**, eabf1444
 84. Li, G., Luan, C., Dong, Y., Xie, Y., Zentz, S. C., Zelt, R., et al. (2021) ExpressHeart: web portal to visualize transcriptome profiles of non-cardiomyocyte cells. *Int. J. Mol. Sci.* **22**, 8943
 85. Ma, W. F., Turner, A. W., Gancayco, C., Wong, D., Song, Y., Mosquera, J. V., et al. (2022) PlaqView 2.0: a comprehensive web portal for cardiovascular single-cell genomics. *Front. Cardiovasc. Med.* **9**, 969421
 86. Pan, H., Xue, C., Auerbach, B. J., Fan, J., Bashore, A. C., Cui, J., et al. (2020) Single-cell genomics reveals a novel cell state during smooth muscle cell phenotypic switching and potential therapeutic targets for atherosclerosis in mouse and human. *Circulation* **142**, 2060–2075
 87. Aran, D., Looney, A. P., Liu, L., Wu, E., Fong, V., Hsu, A., et al. (2019) Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172
 88. Azad, A., Pavlopoulos, G. A., Ouzounis, C. A., Kyrpides, N. C., and Buluç, A. (2018) HipMCL: a high-performance parallel implementation of the Markov clustering algorithm for large-scale networks. *Nucleic Acids Res.* **46**, e33