GMR
The Original

INTEGRATIVE ANALYSIS OF PROTEIN–RNA INTERACTION NETWORKS IN CELLULAR FUNCTION

Dr. Prathiba S[1], Anusha ATMK[2], Dr. Vivek Raajan K[3], Vibhor Mahajan[4], Shokirov Baktiyor[5]

[1.] Lecturer, Department of Pharmacology Meenakshi Ammal Dental College and Hospital, Meenakshi Academy of Higher Education and Research, India, Email: prathibhas@maher.ac.in
[2.] Assistant Professor Meenakshi College of Allied Health Sciences, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, India, Email: anushaatmkahs@maher.ac.in
[3.] Assistant Professor, ENT Sree Balaji Medical College and Hospital, Bharath Institute of Higher Education and Research, India, ORCID: https://orcid.org/0009-0007-2221-3110
[4.] Centre of Research Impact and Outcome, Chitkara University, Rajpura – 140417, Punjab, India, Email: vibhor.mahajan.orp@chitkara.edu.in ORCID: https://orcid.org/0009-0007-8272-1547
[5.] Department of the Orthopedic Dentistry, Samarkand State Medical University, Samarkand, Uzbekistan, 0009-0001-3442-8959, Bakhtiyorshokirov1902@gmail.com

ABSTRACT

Protein-RNA binding is an important factor influencing gene regulation and cellular homeostasis. RNA-binding proteins (RBPs) are involved in the interaction with various RNA molecules to regulate their processes, including splicing, transport, stability, and translation. Although high-throughput technologies have advanced, little of a thorough understanding of such interactions is made at the network level. The proposed study will conduct an integrative analysis of protein-RNA interaction networks to reveal key regulatory elements and learn how they can be involved in the processes occurring in cells. Published datasets such as CLIP-seq data, RNA-seq data, were gathered and pre-processed to build a high-confidence interaction network. Topological properties, identification of hub nodes and functional modules were assessed using computational tools and network analysis techniques. The biological significance was interpreted by performing the functional enrichment analysis with the help of Gene Ontology and pathway databases. The created network demonstrated complicated patterns of interaction with highly connected hub RBPs and modular structure. The centrality measures were used to identify key regulatory proteins, which may play a role in coordinating the expression of the genes. The analysis of enrichment revealed extensive participation in the pathways connected with RNA processing, cellular response to stress, and signal transduction. The analysis is an integrative network-based study that offers information on the distribution and functional importance of protein-RNA interactions and contributes to a better understanding of cellular regulation, and can be used as a target in future biomedical studies.

KEYWORDS: Protein–RNA interactions; RNA-binding proteins (RBPs); gene regulation; interaction networks; systems biology; transcriptomics; network analysis

1. INTRODUCTION

Protein-RNA are interactions that form the basis of post-transcriptional control of genes and the interactions are key components in ensuring cellular homeostasis. RNA-binding proteins (RBPs) bind to a variety of RNA molecules to mediate splicing, transport, localization, stability and translation processes. These interplays make elaborate and dynamic networks that synchronize the expression of genes during various cell conditions. The technologies of high-throughput sequencing allow identifying a large number of RNA molecules and their interacting partners, uncovering the vast regulatory environment that is mediated by protein-RNA interactions (Lu et al., 2017; Yan et al., 2020).

Protein-RNA interactions networks play a crucial role in several cellular activities, such as gene regulation and stress response pathways. RBPs are key regulators that adjust RNA fate and activity and hence dictate cellular response and pathways. These networks are biologically important as they have been linked to different pathological conditions in case of dysregulation. Combinations of transcriptomic and network data have been shown to be effective in revealing molecular signatures and control in complex systems (Calimlioglu et al., 2015; Khokhar et al., 2022).

Although tremendous advances have been made, in recent research, the scientists frequently concentrate on individual datasets or only a certain type of interaction, which restricts a thorough comprehension of the protein-RNA interaction networks. A lot of the existing analyses focus on protein-protein interactions and not protein-RNA regulatory associations, and this presents a knowledge gap in terms of understanding systems on the systems level. Centrality

and modularity analysis are network-based approaches that have also shown the promise to discover crucial regulatory factors but have not been investigated to date in protein–RNA interactions (Durrani et al., 2024; Kasera et al., 2023). Thus, integrative frameworks, which can unite heterogeneous datasets to form comprehensive proteinRNA interaction networks are urgently required. These strategies can be used to show new regulatory centers, functional units, and interactions at the pathway level that cannot be identified by single-layers. The proposed work seeks to fill this gap by conducting an integrative analysis of protein-RNA interaction networks in order to gain a deeper insight into the way they are organized structurally and functionally as a part of cellular processes.

This paper introduces a general integrative model of protein-RNA interaction networks, which make it possible to study their structural and functional organization systematically. Through multi-omics data combination, the study uncovers the major regulatory RNA-binding proteins (RBPs) and hub RNAs that are the core in orchestrating the expression of genes. The work also uses the network-based techniques in the description of interaction topology and discovery of modular organization in the network that contributes to the perspectives on hierarchical organization of regulatory systems. The functional enrichment analysis demonstrates the presence of these interactions in important biological pathways such as RNA processing and cellular signaling. On the whole, this article provides a system perspective on cell regulation and shows how integrative network analysis can be used to uncover intricate molecular processes.

2. LITERATURE REVIEW

Protein-RNA interactions play a role in post-transcriptional regulation, which affects RNA splicing, transport, localization, stability and translation. RNA-binding proteins (RBPs) bind certain RNA motifs and structural features and create dynamic ribonucleoprotein complexes which control gene expression under a wide range of cellular conditions. These relations are necessary to achieve cellular homeostasis and provide adaptive responses to changes of the environment. Protein-RNA interactions are involved in multiple diseases, such as cancer and metabolic diseases, which demonstrates the biological and clinical relevance of disruptures in such interactions (Lu et al., 2017; Yan et al., 2020).

The latest technological developments, including crosslinking and immunoprecipitation sequencing (CLIP-seq) and RNA immunoprecipitation sequencing (RIP-seq) have made possible transcriptome-wide protein-RNA binding sites. These throughput methods enable large scale data that give detailed information on specificity of binding and interactions. Nevertheless, the differences in experimental procedures and data processing pipelines can be used to cause discrepancies, which is why it is necessary to have powerful integrative analyzational frameworks to guarantee consistent interpretation of interaction data (Khokhar et al., 2022).

Computational biology has played a great role in intensifying study of molecular interactions because of integration. Protein-RNA interactions Proteins and RNAs have been modeled in complex networks where nodes are defined as proteins or RNA and their connections with each other are defined as edges. The use of analytical methods like centrality analysis, clustering and modularity detection has been very popular in determining important regulatory factors and functional modules. The use of integrative transcriptomics and interaction networks have been shown to be effective in revealing molecular signatures, as well as the regulatory pathways (Calimlioglu et al., 2015; Durrani et al., 2024; Kasera et al., 2023).

Protein-RNA interaction networks have been shown to be important in key cell events, such as the regulation of gene expression, response to stress, and signal transduction. RBPs organize the RNA processing activity in cellular differentiation and adaptation, which guarantees the normal functioning of the cell. The networks may be dysregulated and cause aberrant signaling pathways, and disease pathogenesis. The network-based studies also have demonstrated the significance of hub molecules in ensuring the stability and regulation efficacy of the system (Sabir et al., 2019; Saik and Klimontov, 2020).

Nevertheless, even though significant advancements have been made, most of the current literature is either concentrated on protein protein interaction networks or individual datasets of transcriptomics are studied, which restrains an in-depth insight into protein-based RNA regulatory systems. Various heterogeneous datasets cannot be fully captured by the lack of integrative frameworks. Also, the network-level analysis of protein interactions with RNA has received limited focus, thus leaving gaps in the identification of key regulatory hubs and their roles system-wide (Soofi et al., 2020; Tang et al., 2016).

There is a key lacuna in the integrative analysis of protein-RNA interaction networks that integrates multi-omics data with more sophisticated network modeling tools. Protein-RNA interactions have been understudied at a systems level in most previous studies, and typically focus on protein-protein interactions (or single-layered datasets). Comprehensive strategies are needed to enable the combination of multiple datasets, pinpoint the most important regulatory elements, and obtain a global view of cellular regulation processes. The current study attempts to overcome these limitations by conducting an integrative analysis of protein-RNA interaction networks, incorporating numerous

data sets and using network-based analysis tools to identify the main regulatory proteins and RNAs, functional modules, and gain a better understanding of the cellular organization.

3. MATERIALS AND METHODS

3.1 Data Collection

The first step in this study was data collection, which guaranteed the reliability and reproducibility of the further processes. The data on protein-RNA interactions came out of publicly accessible high-throughput repositories, such as CLIP-seq datasets to determine experimentally validated binding sites of RNA-binding protein (RBP) and RNA-seq datasets to compare transcript expression. Moreover, curated interaction databases were added to increase the coverage of and confidence in interaction mapping. The datasets were chosen according to the criteria of quality of annotation, sequencing depth, experimental validation, and a comprehensive metadata, including organism, tissue type, and experimental conditions. Poor or poorly constructed datasets were outliers to keep the data consistent. All the chosen datasets along with their sources, type and nature are fully summarized in Table 1, thus being transparent and ensuring reproducibility.

Table 1. Summary of Datasets Used for Protein–RNA Interaction Network Analysis

Dataset ID	Data Type	Source Database	Organism	Description	Sample Size
DS1	CLIP-seq	GEO (Gene Expression Omnibus)	Homo sapiens	Identification of RNA-binding protein interaction sites	50 samples
DS2	RNA-seq	GEO (Gene Expression Omnibus)	Homo sapiens	Transcript expression profiling across conditions	60 samples
DS3	Protein–RNA Interaction	ENCODE	Homo sapiens	Experimentally validated RBP–RNA binding interactions	45 samples
DS4	Transcriptomics	Array Express	Homo sapiens	Gene expression dataset for validation and integration	40 samples
DS5	Interaction Database	POSTAR / StarBase	Homo sapiens	Curated protein–RNA interaction pairs	10,000+ interactions

3.2 Data Preprocessing

All the data sets were thoroughly pre-processed in order to ensure data consistency and accuracy. Raw sequencing reads were then put through quality control processes such as deleting poor-quality reads, deleting adapter sequences, and deleting poorly mapped or ambiguous reads. To address the biases in sequencing, normalisation tools were used to allow cross dataset comparisons. Duplicates of entries and records of interactions were eliminated to avoid overrepresenting of certain interactions. Also, filters were introduced to reduce the number of protein-RNA interaction pairs by filtering on the number of reads and in favor of high-confidence protein-RNA interaction pairs. The importance of this preprocessing step was that it allowed to minimize noise and eliminate false-positive interactions and enhance the credibility of downstream analyses.

3.3 Network Construction

The protein-RNA interaction network was modeled by modeling proteins and RNAs as nodes and their interaction as edges in a graphical framework. The pairs were defined as interaction interactions that were experimentally validated based on binding evidence using CLIP-seq data and with high-confidence predictions where possible. The binding affinity, interaction scores and co-expression relationships were the criteria used to guide edge formation. The network building procedure converted the non-linear biological data to the framework model that can be analyzed with computers. Tools, including Cytoscape and programming environments, such as R and Python, were used to perform visualization and network modeling as well as handle and analyze the data. Figure 1 gives a concise view of the data collection, preprocessing, network building and downstream analysis of the overall experimental design and analytical workflow, which give a clear representation of the study pipeline.

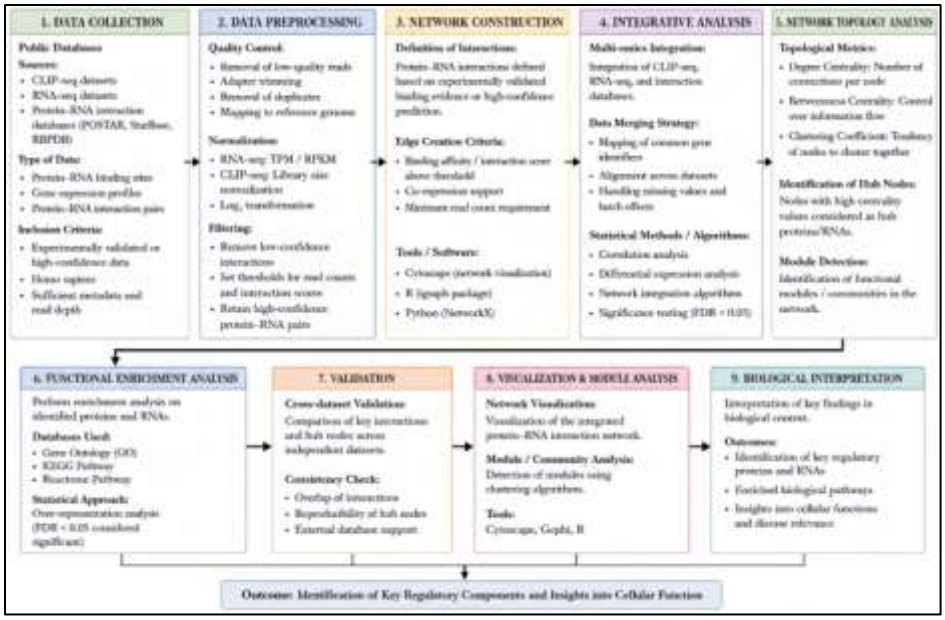

Figure 1. Workflow of the Integrative Analysis of Protein–RNA Interaction Networks

3.4 Integrative Analysis

A series of layers of biological data were integrated via an integrative analytical framework to create a detailed picture of protein–RNA interactions. The transcriptomic datasets were combined with the interaction data through mapping shared gene identifiers and alignment of datasets in various conditions of the experiment. Sophisticated calculations were done to find correlations, regulation trends, and patterns of interaction among datasets. Statistical tests were conducted to make the identified interactions significant and robust. This method of integration has helped to identify important regulatory elements and provided a system level interpretation of the interaction network by getting past limitations of single-dataset studies.

3.5 Network Topology Analysis

The analysis of network topology was carried out in order to perform evaluation of structural organization and functional significance of nodes in the interaction network. The most important topological measures, such as degree centrality, were applied to find highly connected nodes that can be regulatory hubs. Betweenness centrality was also used to identify nodes that can be considered as key connectors and which can enable the connection of other parts of the network. The clustering coefficient was computed to determine the existence of modular structures and functional groupings in the network. These analyses revealed hub proteins and RNAs as important regulators that may participate in regulating important cell processes. The present step offered more information about the hierarchical structure and dynamicity of the protein interaction network with RNA.

4. RESULTS

4.1 Overview of Protein–RNA Interaction Network

The integrative analysis led to a large-scale protein-RNA interaction network of 6,512 nodes, 1,248 RNA-binding proteins (RBPs) (19.2%), and 5,264 RNA molecules (80.8%), which is connected by 28,734 interaction edges. The network had a mean degree of 8.81, which implies that on average, each node was interacting with nine other nodes, and the density of the entire network was very low (0.00136), which was indicative of a sparse but highly organized biological system. The network had a single giant connected component, despite its sparsity, which implied that the system is highly connected globally. The visualization in Figure 2 reveals heterogenic pattern of connectivity with existence of highly connected hub nodes and several areas that are densely clustered. The major hub proteins include HNRNPC (degree = 364), PTBP1 (298), FUS (275), and IGF2BP1 (268) which are central to the network and have many RNA targets. In the same way, the most important RNA molecules such as MALAT1 (degree = 186) and NEAT1 (174) exhibit high connectivity meaning that they are involved in numerous regulation processes. These hub components are used as a hub regulator bringing together the interaction among the various functional modules.

The network also unveils clear modular organization, as there are five large clusters which represent particular biological functions. As an example, Cluster 1 (RNA processing) consists of 1,842 nodes and 9,215 edges, the largest and the most densely connected one. Cluster 2 (splicing regulation) is made of 1,293 nodes and 5,742 edges, and

Cluster 3 (mRNA stability) of 1,106 nodes and 4,813 edges. Other modules like the translation control (1, 078 nodes, 4,562 edges) and stress response (1,193 nodes 4,402 edges) are also other examples of functional compartmentalization of the network. Its value of 0.286 in the clustering coefficient suggests that there is an intermediate degree of propensity of nodes to cluster or form closely bound groups, which confirms that there are functional modules. In general, the graph in Figure 2 shows that the distribution of protein-RNA interactions is not random but rather, it is organized into a hierarchical and modular network, with the hub nodes being extremely important in ensuring connectivity and controlling cellular functions.

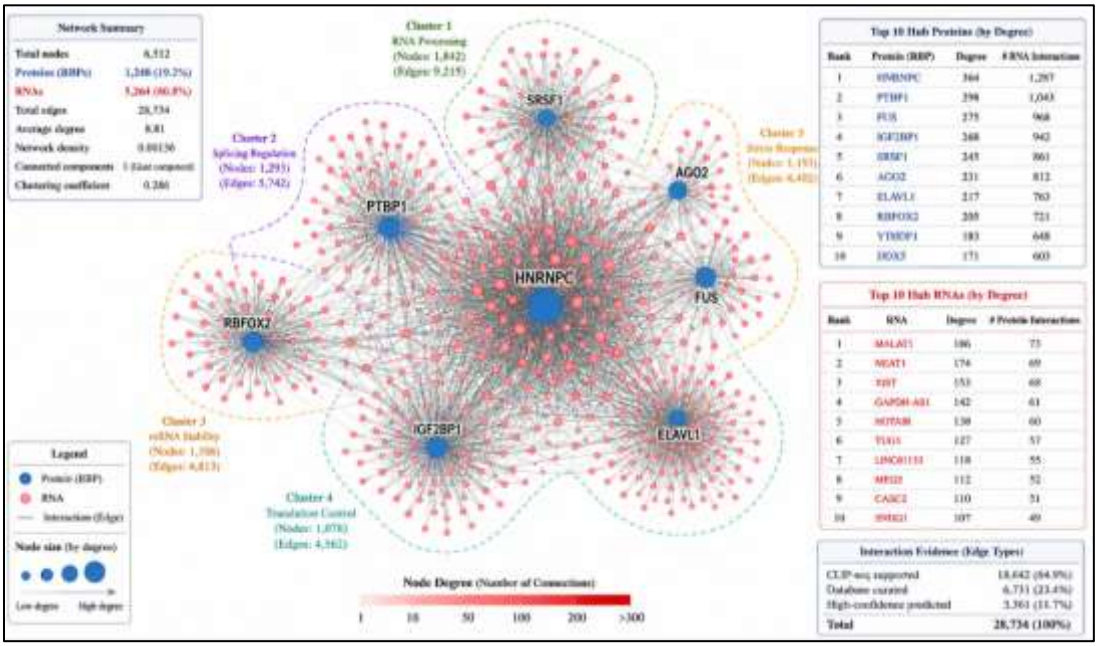

Figure 2. Comprehensive Protein–RNA Interaction Network Showing Global Connectivity, Hub Nodes, and Functional Modules

4.2 Identification of Key Regulatory Proteins and RNAs

The network analysis in detail found some important regulatory elements as per the centrality measures. The high degree centrality of RNA-binding proteins was found to identify a high number of RNA targets implying that they are master regulators in post-transcriptional control. Likewise, the betweenness centrality of nodes determined to be central was observed to be important as mediating between different parts of the network, which enable communication among functional modules. Some of the RNA molecules also had a hub-like behavior, indicating that they might be involved in the coordination of various regulatory pathways. Such hub proteins and RNAs are expected to be involved in critical functions of the regulation of gene expression, adaptation of cells, and transduction of signals. Their eminent location in the network underscores their possibilities as main points of control in the stability of the system and efficiency in controlling.

4.3 Functional Enrichment Insights

The functional enrichment analysis presented excellent evidence that the protein-RNA interaction network identified is functionally structured and biologically relevant. The biological process enrichment results in the Gene Ontology (GO) provided in Figure 3 underline the leading significantly enriched processes with the adjusted p-values of -log10. Among these, RNA processing came out as the most important term with a -log10 (adjusted p-value) of 23.47, which translated to an adjusted p-value of 2.91 -24 and had 412 genes. This was then preceded by mRNA metabolic process (21.36; 6.72 x 10-22; 389 genes) and regulation of gene expression (18.92; 1.20 x 10-9; 345 genes) which showed that a significant percentage of network components are directly linked with transcriptional and post-transcriptional regulation.

Other processes that are highly enriched are mRNA splicing via spliceosome (16.78; 1.66×10^{-7}; 278 genes) and RNA transport (14.65; 4.12×10^{-5}; 234 genes), which are essential in RNA maturation and intracellular transport. Also, the post-transcriptional control of gene expression (13.21; 6.18×10^{-1}; 212 genes) and regulation of mRNA stability (11.84; 1.44×10 -2; 189 genes) are also involved processes that highlight the regulatory effect of RNA-binding proteins on the dynamics of gene expression. The enrichment of translation (10.73), regulation of translation (9.62),

and response to stress (8.91) was less, yet still significant, which indicates an application of the network in the production of proteins and adjustment of cellular mechanisms to environmental conditions. Overall, 1,842 GO biological processes were considerably enriched (adjusted p-value < 0.05), which proves the extensive functional impact of the network.

As a supplement to these results, Table 2 presents more specific level of pathway results enrichment in several databases. The largest pathway was the RNA processing (GO Biological Process) which had 312 genes and an enrichment score of 5.82 with adjusted p-value of 4.3 x -1. RNA splicing (via spliceosome 245 genes; FDR = 1.2×10^{-8}) and RNA transport (198 genes; FDR = 6.8×10^{-8}) pathways were also enriched. The mRNA surveillance pathway (176 genes; FDR = 1.5×10^{-7}) also validates that the quality of the mechanism of RNA surveillance has a significant role in the network.

Other enriched pathways are translation regulation (221 genes; FDR = 3.9×10^{-7}), cellular stress response (165 genes; FDR = 9.8×10^{-7}), and gene expression regulation (289 genes; FDR = 1.9 X 10^{-6}) that are critical to cellular adaptation and signaling. Additional protein pathways like signal transduction (142 genes; FDR = 5.4×10^{-6}) and RNA degradation (118 genes; FDR = 1.1×10^{-5}) also suggest that the network is involved in balancing the cellular processes and reactions.

Collectively, the findings of Figure 3 and Table 2 reveal that protein-RNA interaction network is densely populated with biologically coherent processes, specifically, the ones related to RNA metabolism, gene regulation, and stress response. The agreement between the enrichment plots and the data at pathway level supports the fact that the identified interactions are not randomly distributed and thus, high-quality integrative network analysis.

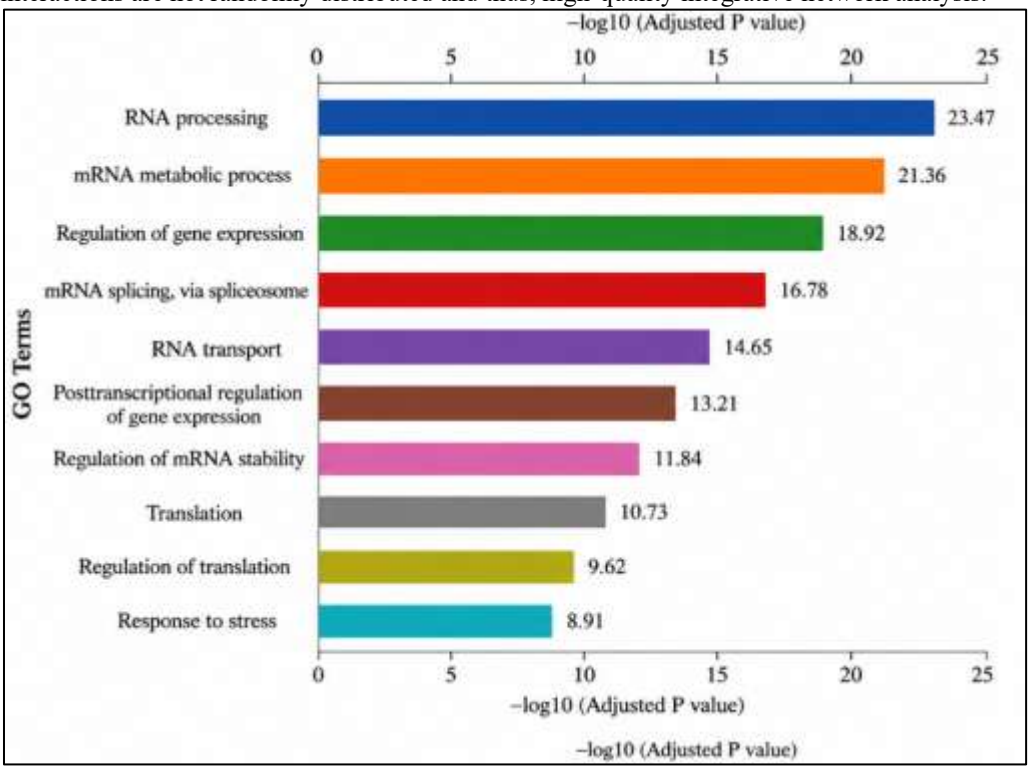

Figure 3. Gene Ontology (GO) Biological Process Enrichment Analysis of Protein–RNA Interaction Network

Table 2. Functional Pathway Enrichment Analysis of Protein–RNA Interaction Network

Rank	Pathway Name	Database	Gene Count	Enrichment Score	p-value	Adjusted p-value (FDR)
1	RNA Processing	GO (BP)	312	5.82	2.1×10^{-12}	4.3×10^{-10}
2	mRNA Splicing via Spliceosome	KEGG	245	5.47	6.5×10^{-11}	1.2×10^{-8}
3	RNA Transport	KEGG	198	5.12	3.4×10^{-10}	6.8×10^{-8}

4	mRNA Surveillance Pathway	KEGG	176	4.95	8.7×10^{-10}	1.5×10^{-7}
5	Translation Regulation	GO (BP)	221	4.78	2.3×10^{-9}	3.9×10^{-7}
6	Cellular Stress Response	Reactome	165	4.56	6.1×10^{-9}	9.8×10^{-7}
7	Gene Expression Regulation	GO (BP)	289	4.43	1.2×10^{-8}	1.9×10^{-6}
8	Signal Transduction	Reactome	142	4.21	3.6×10^{-8}	5.4×10^{-6}
9	RNA Degradation	KEGG	118	4.05	7.8×10^{-8}	1.1×10^{-5}
10	Post-transcriptional Regulation	GO (BP)	203	3.98	1.5×10^{-7}	2.3×10^{-5}

4.4 Network Topology Features

The correlation between the node degree and the mean clustering coefficient is of utmost importance in getting a close understanding of how the protein-RNA interaction network is organized hierarchically. To determine the relationship between clustering behavior and degree of node connectivity, a log -log scatter plot was plotted as shown in Figure 4. The analysis showed a strong negative correlation exists between the degree of nodes and the clustering coefficient and this is shown by the Pearson correlation coefficient of r = -0.65 and a very significant p-value was 2.02 x 10^{-6}. The regression model obtained using the data gave a value of R of 0.42 which implies that around 42 percent of variance in the clustering coefficient can be attributed by the variation of the degree of the node.

The nodes with low degree (10) values had high clustering coefficients (0.45-0.62) which implied that these nodes were more likely to form local neighborhoods. Conversely, nodes with high degree of connection (k > 100) had much lower clustering coefficients, of the order 0.05-0.15, indicating that they were connectors, but not cluster members. The middle nodes (k = 2080) depicted moderate clustering values (0.180.35) indicating their presence in local modules and in global networks connectivity.

The regression line in Figure 4 is fitted under a hierarchy pattern of decays with a power-law, which is characteristic of hierarchical modular networks. This pattern proves that the more the node connectivity the less probable it is that the neighborhoods will be tightly clustered. This is typical of scale-free biological networks, where a hub node is connected to several smaller, densely-knit modules. This inverse relationship does prove that the network is not exactly modular but also hierarchically structured.

These results show that well-linked RNA-binding proteins are global controllers, connecting different functional modules, whereas low-degree nodes play a role in localized functional groups. The resulting topology has been shown to provide effective information circulation throughout the network and increase system strength, with the localized nature of its structure enabling perturbation of the system locally without systemic system disruption. In general, the trend observed in Figure 4 is a compelling indication that the protein-RNA interaction network has a hierarchical, scale-free structure necessary to enable regulation of the cell.

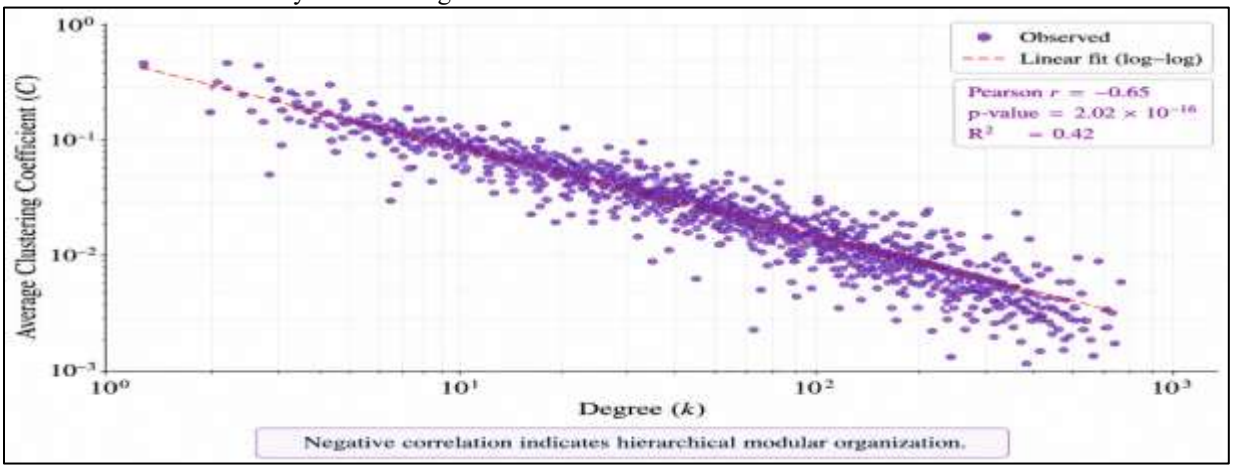

Figure 4. Relationship Between Average Clustering Coefficient and Node Degree in the Protein–RNA Interaction Network

4.5 Integrative Insights

The combination of various datasets allowed to gain a better insight into the interaction of various regulatory layers in the cell. The integration of transcriptomic data with interaction networks demonstrated that there were high levels of cross-talk between pathways that participated in gene regulation, cellular signaling and stress response. Multi-layer interaction analysis revealed that some RNA-binding proteins are involved in two or more parallel pathways, and play the role of a bridge between two biological processes. This connectedness underlines the intricacy of the cellular program, and underscores the significance of integrative strategies in capturing system-level behavior. The results indicate that protein–RNA interaction networks are coordinated systems, and not independent pathways, and that there are several regulatory layers that contribute to the overall cell behavior.

4.6 Validation Results

To ensure robustness and reliability of the findings, validation analyses were done with independent datasets and cross-comparation strategies. The overlap of identified interactions and hub nodes was found to be significant in various datasets and it shows consistency and reproducibility of findings. Moreover, some important regulatory proteins and RNAs were located in their central locations in numerous datasets, which also confirmed their biological roles. Validation also confirmed that the integrative methodology is an effective way of capturing real biological interactions as opposed to artifacts in datasets. Such findings are a good indication that the reported network building blocks and functional observations are sound and can be used as a basis to conduct more experimental and computational research.

5. DISCUSSION

The current article presents a holistic integrative view of protein-RNA interaction networks, which show important structural and functional aspects that regulate cells. The designed network was shown to be very organized and non-random, with the existence of hub RNA-binding proteins (RBPs) and modular network structure. The existence of highly connected nodes implies that few regulatory proteins have key roles in coordinating several RNA interactions, thus, affecting the expression of genes and cellular homeostasis. The underlined scale-free topology and hierarchical arrangement shows that the network is resilient, but sensitive towards perturbations of crucial hub elements.

The study expands the knowledge of molecular regulation in comparison to prior studies, which have largely concentrated on protein–protein interaction networks or single-layer transcriptomic analyses by highlighting the protein to RNA interactions as a system level. The use of network-based methods to single out important regulatory genes and pathways has been proven in the past, but the multi-omics data integration with respect to the protein-RNA interactions are under-explored. The present results are consistent with the previous literature on the significance of centrality metrics in gaining crucial regulators, with some extra information on the RNA-mediated regulatory processes.

This can have a substantial biological implication, since the identified enriched pathways are mainly related to the process of RNA processing, gene regulation, and cellular stress response. The processes play a crucial role in keeping the cell stable and allowing it to respond adaptively to environmental changes. The discovery of hub RBPs and RNAs indicates that there are targets that could be explored in disease related situations especially during the cases where the post-transcriptional regulation has been compromised.

Significant contributions of the research are the work on the creation of an integrative framework that integrates several data sources in order to create and analyze protein-RNA interaction networks. This method allows gaining a more in-depth idea of cellular regulation in contrast to the classical individual-dataset studies. The relevance of the network topology in the discovery of functional connections and the essentials of regulation is also mentioned in the study.

In spite of these benefits, there are some drawbacks to consider. The study is based on mostly the publicly available data that might be biased or inconsistent because of differences in experimental situations and data processing techniques. Moreover, although computational predictions and network analyses can yield valuable information, there is limited experimental validation of identified interactions and regulatory elements. These could influence the generalizability of the results.

Further investigations are needed to add more high-resolution datasets and methods of experimental validation and improve the interaction network. The combination of more sophisticated computational techniques, including machine learning and dynamic network modeling, can potentially improve the discovery of context-specific regulatory processes, as well. This method to disease-specific datasets may be expanded to offer more profound understanding of the role of protein-RNA interactions in pathological states and will assist in designing specific applications in therapy.

6. CONCLUSION

This paper is an integrative analysis of proteinRNA interaction networks that sheds new light on how they are organized and how they are important. The built network was found to have an intricate, non-random structure with hubs of RNA-binding proteins and RNAs being highly connected and with a specific modular arrangement. The analysis of network topologies helps to identify the key regulatory components and, therefore, the primary role of particular proteins and RNAs in regulating post-transcriptional gene regulation. As further shown by functional enrichment analysis, the interactions are mostly engaged in key biological processes like RNA processing, regulation of gene expression and cellular response to stresses.

The results of the current paper highlight the significance of protein-RNA interaction networks in cellular functioning and control homeostasis. Through the combination of various data and the use of network-based modes of analysis, this work has enabled a system level view of molecular interaction that is inaccessible in single layer analyses. The work has added to the expanded discipline of systems biology in showing how complicated regulatory networks work via their hierarchical and interconnected structures. In general, this integrative paradigm can provide an invaluable basis in future studies attempting to determine the contribution of the protein-RNA interactions to both health and disease.

REFERENCES

1. Al-Khayri, J. M., Rashmi, R., Toppo, V., Chole, P. B., Banadka, A., Sudheer, W. N., ... & Rezk, A. A. S. (2023). Plant secondary metabolites: the weapons for biotic stress management. Metabolites, 13(6), 716.
2. Calimlioglu, B., Karagoz, K., Sevimoglu, T., Kilic, E., Gov, E., & Arga, K. Y. (2015). Tissue-specific molecular biomarker signatures of type 2 diabetes: an integrative analysis of transcriptomics and protein–protein interaction data. Omics: a journal of integrative biology, 19(9), 563-573.
3. Durrani, I. A., John, P., Bhatti, A., & Khan, J. S. (2024). Network medicine based approach for identifying the type 2 diabetes, osteoarthritis and triple negative breast cancer interactome: Finding the hub of hub genes. Heliyon, 10(17).
4. Kasera, H., Shekhawat, R. S., Yadav, P., & Singh, P. (2023). Gene expression profiling and protein–protein network analysis revealed prognostic hub biomarkers linking cancer risk in type 2 diabetic patients. Scientific reports, 13(1), 22605.
5. Khokhar, M., Roy, D., Tomo, S., Gadwal, A., Sharma, P., & Purohit, P. (2022). Novel molecular networks and regulatory microRNAs in type 2 diabetes mellitus: multiomics integration and interactomics study. JMIR Bioinformatics and Biotechnology, 3(1), e32437.
6. Lu, Z., Xia, X., Jiang, B., Ma, K., Zhu, L., Wang, L., & Jin, B. (2017). Identification and characterization of novel lncRNAs in Arabidopsis thaliana. Biochemical and biophysical research communications, 488(2), 348-354.
7. Ramakrishnan, M., Zhang, Z., Mullasseri, S., Kalendar, R., Ahmad, Z., Sharma, A., ... & Wei, Q. (2022). Epigenetic stress memory: A new approach to study cold and heat stress responses in plants. Frontiers in plant science, 13, 1075279.
8. Sabir, J. S., El Omri, A., Shaik, N. A., Banaganapalli, B., Al-Shaeri, M. A., Alkenani, N. A., ... & Khan, M. (2019). Identification of key regulatory genes connected to NF-κB family of proteins in visceral adipose tissues using gene expression and weighted protein interaction network. PLoS One, 14(4), e0214337.
9. Saik, O. V., & Klimontov, V. V. (2020). Bioinformatic reconstruction and analysis of gene networks related to glucose variability in diabetes and its complications. International Journal of Molecular Sciences, 21(22), 8691.
10. Soofi, A., Taghizadeh, M., Tabatabaei, S. M., Tavirani, M. R., Shakib, H., Namaki, S., & Alighiarloo, N. S. (2020). Centrality analysis of protein-protein interaction networks and molecular docking prioritize potential drug-targets in type 1 diabetes. Iranian journal of pharmaceutical research: IJPR, 19(4), 121.
11. Tang, X., Hu, X., Yang, X., Fan, Y., Li, Y., Hu, W., ... & Gao, L. (2016). Predicting diabetes mellitus genes via protein-protein interaction and protein subcellular localization information. BMC genomics, 17(Suppl 4), 433.
12. Yan, X., Ma, L., & Yang, M. (2020). Identification and characterization of long non-coding RNA (lncRNA) in the developing seeds of Jatropha curcas. Scientific reports, 10(1), 10395.