



Exploring lung cancer protein network: Understanding structure and function through metric space modeling

Emad Fadhal

*Department of Mathematics Statistics, College of Science, King Faisal University,
P. O. Box 400, Al-Ahsa 31982, Saudi Arabia*

Abstract. Lung cancer remains a significant health threat with high mortality rates. Using graph theory, we modeled the protein-protein interaction (PPI) network in lung cancer to explore its complex structure. This approach allows for the analysis of network properties and the identification of key proteins driving biological processes. Our analysis revealed RPS27A as a central protein within the network, associated with diverse functions related to ribosome biogenesis, translation, cell growth, apoptosis, and cancer progression. This suggests that RPS27A may have multiple functions in cancer development and progression, including in MAPK signaling pathways. Importantly, our study uniquely identifies RPS27A as a central hub in lung cancer PPI networks, shedding light on its pivotal role in disease pathogenesis. Additionally, we identified a central network zone enriched with proteins involved in key signaling pathways, presenting novel insights into potential therapeutic targets for lung cancer treatment. Pathway enrichment analysis further highlighted functional specialization across network zones, providing a comprehensive understanding of the intricate interplay between biological pathways in lung cancer progression. This study underscores the multifaceted roles of central proteins like RPS27A within lung cancer's PPI network and the network's potential for pinpointing therapeutic targets, presenting a novel perspective on the intricate network of molecular interactions driving lung cancer pathogenesis.

2020 Mathematics Subject Classifications: 05C12, 92-08

Key Words and Phrases: Protein interaction networks, metric spaces, lung cancer, molecular mechanisms, therapeutic targets

1. Introduction

Lung cancer remains one of the most prevalent and lethal malignancies worldwide, posing a significant public health challenge [30]. Despite advancements in understanding its intricate molecular mechanisms [22], identifying effective therapeutic targets remains crucial. While PPI networks are invaluable in uncovering potential targets [1] [46] [12] [24] [8]. PPI networks have emerged as a valuable tool. These networks map the intricate web of interactions between proteins within a cell, providing insights into key signaling

DOI: <https://doi.org/10.29020/nybg.ejpam.v17i2.5075>

Email address: efadhal@kfu.edu.sa (E. Fadhal)

pathways and potential therapeutic targets. Their construction, facilitated by both experimental and computational methods, allows researchers to visualize the complex interplay of proteins that dictate cellular activities [27].

The application of PPI networks in lung cancer research has yielded remarkable insights. These networks have aided in identifying pivotal "hub proteins" like EGFR and TP53, known to play critical roles in tumorigenesis and disease progression [46]. Additionally, PPI networks have proven instrumental in discovering novel biomarkers for early diagnosis and improved patient prognosis [12] [24]. Furthermore, it is vital to identify crucial proteins through PPI analysis to gain a comprehensive understanding of cellular processes, disease mechanisms, and potential targets for therapeutic [7]. Current approaches often lack a unified framework, hindering their full potential. Additionally, while graph theory finds broad application in modeling molecular structures [44] [32] [10] [45] [33], its integration with PPI network analysis in lung cancer research is limited [31].

Exploring PPIs in the context of lung cancer research has become a cornerstone in deciphering the intricate molecular mechanisms of this deadly disease. By revealing key interaction partners and providing insights into both structural and functional aspects, PPI networks offer a multifaceted approach to advancing our understanding of lung cancer biology. This study focuses on the PPI network in lung cancer, aiming to identify key proteins and pathways that could serve as potential therapeutic targets. We established and modeled the lung cancer protein-protein interaction network (LCPIN) as a metric space. This approach leverages the well-defined properties of topological spaces, drawing upon a rich history and extensive theoretical foundation [9] [4]. This strategic framework enables a comprehensive spatial perspective on PPIs, utilizing distance as the primary metric. While protein weighting is not employed, the method precisely pinpoints the relative positions of nodes, even in vast networks containing hundreds of thousands of proteins and numerous interactions. We formally defined the network's center(s) as "the protein(s) with the minimal maximal distance to others" and subsequently categorized all proteins into zones based on their distance from the center. This allows us to pinpoint the exact location of any protein relative to the centre and its neighboring proteins within the network swiftly. This effort aims to provide valuable insights into both network structure and function. Through the elucidation of complex protein interactions and the identification of key players in tumorigenesis, these networks provide valuable insights into disease mechanisms and potential therapeutic targets.

2. Materials and Methods

2.1. Generation of cancer graphs

Our method to constructing a crossover network of proteins associated with cancer involves the mapping proteins expressed in tumors onto the human interactome. We accomplish this by leveraging interactions within the human protein network, and employing the subsequent algorithm.

- (i) Construct a graph G over the binary PPI list of human protein network.

- (ii) Identify tumor proteins that are consistently expressed in cancers in list L ;
- (iii) If the proteins are included in list L , construct the induced interaction data from G .
- (iv) Output-induced cancer graph G' Figure 1.

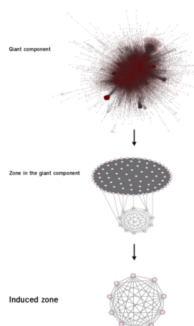


Figure 1: Diagram illustrating the analysis of the giant component of lung cancer for the induced zone.

2.2. Evaluation of lung cancer PPI as metric spaces

In this work, we model PPI networks through the lens of graph theory. A graph, denoted $G = (V, E)$, consists of a set of nodes (V) and a set of edges (E) connecting these nodes. Importantly, the relation E is not reflexive (no self-loops) and symmetric (if x and y are connected, then y and x are connected as well). Nodes represent proteins, and edges represent interactions between them. We interchangeably use "protein" and "node" for clarity. Within the context of the organisms studied, we disregard reflexivity (proteins interacting with themselves) to avoid redundancy. Notably, reflexivity does not influence distance-related concepts. The order of a graph signifies the number of nodes, while the size represents the total number of interactions. A graph is considered complete if every node is connected to all others. When a node x is connected to y , we say y is adjacent to x , denoted by xy . The neighborhood of x encompasses all nodes directly connected to it. The degree of a node (protein) reflects the number of proteins it interacts with in the network. A subgraph of G is a graph where the set of nodes is a subset of V , and the edges are a subset of E . An **induced subgraph** H of G inherits all edges present in G between its nodes. Conversely, if H lacks any edges defined in G , it is not induced. Notably, induced subgraphs are crucial in PINs for defining pathways and processes. A path in a graph refers to a sequence of distinct nodes (v_0, v_1, \dots, v_k) where consecutive nodes are connected by edges. The length of a path is defined as k . A graph is considered connected if a path exists between any two nodes. Otherwise, it is disconnected. Furthermore, components are maximally connected subgraphs. A giant component comprises a majority of the nodes within the entire graph. The distance between two nodes represents the length of the shortest path connecting them. Together, the graph and its corresponding distances define a **metric space**. The eccentricity of

a node v signifies the length of the longest path connecting it to all other nodes. The diameter of the graph is the highest eccentricity among all nodes, while the nodes with the lowest eccentricity are considered the centre of the graph. To evaluate the LCPIN, we conceptualized it as metric spaces and examined the node distances through graph theory. A distinctive strategy to achieve this was by utilizing a Python wrapper for the C++ Boost Graph Library (<http://www.boost.org/>) and executing the Dijkstra algorithm. This algorithm was applied to compute the shortest distances among protein pairs within the network. The proteins showing the least maximum distance to adjacent nodes were recognized as the network's central point or points. Through this approach, the nodes were classified and divided into zones based on their distance to the central point Figure 2.

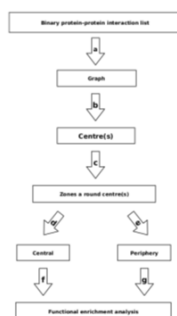


Figure 2: Diagram illustrating the analysis of Protein-Protein Interactions (PPIs) as metric spaces, including: (i) Representation of interactions as a graph model; (ii) Identification of central proteins; (iii) Categorization of proteins into zones; (iv) Highly connected proteins; (v) Less connected proteins; (vi) Proteins of critical importance, often associated with housekeeping functions; and (vii) Proteins exhibiting a diverse range of routine metabolic functions.

2.3. PPI data sources

We consider an interaction network of human functional proteins composed of 9448 nodes and 181706 connections [41].

2.4. Datasets pertaining to gene expression in cancer

The lung cancer dataset we analyzed was obtained from the Gene Expression Bar-code. We utilized the database available at (<https://www.hsls.pitt.edu>). To ensure consistency, we specifically chose the "unified tissue" selection and set the threshold to 0.99. For our analysis, we employed the Human HGU133 platforms.

2.5. Enrichment analysis for pathways and functions

To identify the biological relevance of zones in the LCPIN, proteins were classified into groups according to their closeness to the centre. An over-representation pathway analysis was performed on the groups of proteins linked with each zone, to identify if any

specific functions were attributed to those zones. To perform an enrichment analysis of zones, we used the gene set enricher web service in comparative toxigenomics databases (<http://ctdbase.org/tools/enricher.go>), (<https://david.ncifcrf.gov/>) as well as enrichment analysis of gene ontology terms (<http://www.geneontology.org/GO.tools>) and a significance level of 0.01 was selected as the threshold for statistical significance. Finally, to determine whether zones exhibit functional specialization, we computed the ratio of proteins participating in each enriched pathway.

2.6. Evaluation of the score and core pathways of proteins are oncogene and tumor suppressor

We evaluated the score (the interactions that received high scores are enriched in genes that have been proven to cause cancer through mutations) of proteins that expressed in oncogene and suppressor and their core pathways from Genome-wide sequencing studies of cancer [34].

3. Results

3.1. LCPIN is analyzed using a metric space with a dense core and sparse surrounding structure

To build the network, we modeled LCPIN as a metric space. The goal is to categorize proteins into different zones using their graph-based measure of distance from the central protein while identifying the topological centre. For example, the first zone is the connection with the protein one distance away from the centre of the topology; the second zone is two distances away; and so on Figure 3. The LCPIN we examined has 409 proteins and 4473 interactions. Through modeling, we found that the RPS27A protein is at the centre, which has demonstrated involvement in cancer. Several studies have indicated that RPS27A is overexpressed in different forms of cancer and is involved in regulating cell proliferation, apoptosis, and tumor growth. It has also been associated with drug resistance and a poor prognosis in certain cancers [19] [37] [47].

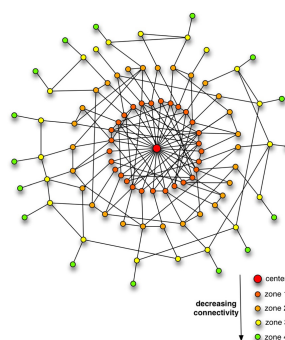


Figure 3: Model capturing the spatial arrangement of PINs based on their distance from a central point.

A summary of the distribution of nodes and the average degree from the centre of LCPIN is shown in Table 1.

Table 1: Summary of the distribution of nodes and average degree from the centre of LCPIN.

Zone	Number of proteins	Average degree	Highest degree	Lowest degree
Centre	1	100%	105	0
1	105	60.9%	84	1
2	123	12.5%	41	1
3	111	4.5%	15	1
4	28	4.9%	20	1
5	21	10.9%	17	1

3.2. Distribution of crucial, signalling, growth-related, cell cycle-regulating, tumor suppressor, oncogenic and therapeutic target proteins within BCPIN.

Table 2 shows the distribution of essential proteins in the LCPIN zone. To assess the LCPIN zone, we selected a list of human proteins that may be more important for orthologous knockout in mice [5]. As shown below, we find that the highest proportion of these proteins is found near the centre. Zones 1-4 have the highest percentage of essential proteins, with 10.48%, 32.52%, 24.32%, 25, and 0% of the proteins in zones 1, 2, 3, 4, and 5, in order.

Table 2: Distribution of essential proteins in zones of BCPIN.

Zone	Number of proteins	Essential	Total
1	1	11(10.48%)	85
2	105	40(32.52%)	
3	123	27(24.32%)	
4	111	7(25%)	
5	28	0(0%)	

Table 3 shows the distribution of signaling proteins in the LCPIN zone. Again, the same observation was made for signaling proteins. They account for 20.95%, 32.52%, 46.84%, 25%, and 9.52% of the proteins present in zones 1, 2, 3, 4, and 5, respectively. In addition, the centre was also determined as a protein that functions in signaling. The signaling pathway has long been considered an attractive avenue for cancer therapy. The authors of [3] showed that protein kinases are a class of enzymes that contribute significantly to regulating cell functions and controlling cell proliferation. Furthermore, this protein kinase cascade has been shown to have particular utility in the treatment of cancer [29]. This offers additional evidence for the potential importance of proteins in the central zone for drug target discovery.

Table 3: Distribution of essential proteins in zones of BCPIN.

Zone	Number of proteins	Signalling	Total
1	1	22(20.95%)	123
2	105	40(32.52%)	
3	123	52(46.84%)	
4	111	7(25%)	
5	28	2(9.52%)	

The distribution of growth proteins in the LCPIN zone is shown in Table 4. Experimental evidence substantiates the involvement of the insulin-like growth factor (IGF) signaling system in cancer progression, persistence, and treatment [11]. A small number of proteins involved in growth function were found to be located in the central zone. These include 0.95%, 2.43%, 2.70%, 0%, and 0% of zones 1-5, with proteins distributed accordingly. This suggests that the central location zone may potentially be a good drug target.

Table 4: Distribution of growth proteins in zones of BCPIN.

Zone	Number of proteins	Growth	Total
1	1	1(0.95%)	7
2	105	3(2.43%)	
3	123	3(2.70%)	
4	111	0(0%)	
5	28	0(0%)	

Table 5 shows the distribution of cell cycle proteins in the BCPIN zone. During the cell cycle, zones 1 and 2 have the highest percentage of these proteins. They make up 19.51%, 9.85%, 2.19%, 0.80%, and 0 of proteins in zones 1-5, in that order. In [16], the authors showed that established cell cycle kinases have limitations as potential targets for anticancer drug discovery. This study also presented a new approach for the development of therapeutic interventions to suppress tumor development and disease progression. This provides compelling evidence of the importance of zone 1-3 proteins in cancer biology and that many of them may be potential new drug targets.

Table 5: Distribution of cell cycle proteins in zones of BCPIN.

Zone	Number of proteins	Cell cycle	Total
1	1	18(17.14%)	36
2	105	14(11.38%)	
3	123	4(3.60%)	
4	111	0(0%)	
5	28	0(0%)	

The distribution of suppressor proteins in the LCPIN zone is shown in Table 6. Proteins

were exclusively localized in zones 1 and 2, with zones 1 and 2 containing the tumor suppressors at 0.95% and 0.81%, respectively. While zones 3-5 lack tumor suppressors, i.e., 0% exists in all cases.

Table 6: Distribution of suppressors proteins in zones of BCPIN.

Zone	Number of proteins	Suppressors	Total
1	1	1(0.95%)	2
2	105	1(0.81%)	
3	123	0(0%)	
4	111	0(0%)	
5	28	0(0%)	

In Table 7, it is clear that zone 2, which has the highest proportion of proteins expressed by oncogenes, predominates. These contain 0%, 2.43%, 0%, 0%, and 0% of proteins in zones 1-5 respectively. According to a study conducted by the authors of [20], the cascade is from oncogenes and tumor suppressors. It causes the development of metastatic prostate cancer. Furthermore, they have been shown to have essential functions in the development of human cancers [21]. Moreover, targeted therapy of these important tumor suppressors and oncogenes shows promising therapeutic potential [17]. This demonstrates the importance of zone 1-3 proteins in cancer biology and the potential of many of them as potential new drug targets, as they are central to tumorigenesis and disease progression.

Table 7: Distribution of oncogenes proteins in zones of BCPIN.

Zone	Number of proteins	Oncogenes	Total
1	1	0(0%)	3
2	105	3(2.43%)	
3	123	0(0%)	
4	111	0(0%)	
5	28	0(0%)	

Finally, zones 1-4 are enriched for successful drug targets. The distribution of these proteins is shown in Table 8. These proteins account for 0.95%, 2.43%, 2.70%, and 3.57% of the proteins in zones 1-4 respectively. In fact, this provides perhaps the most compelling evidence for the importance of a centrally located zone potentially good for drug targets.

3.3. Distribution of MAPK signalling cascade, positive regulation and negative regulation of signalling, apoptosis positively regulated and negative regulation in lung cancer PIN

Table 10 shows that results related to MAPK signaling cascades, positive and negative regulation of signaling, and positive and negative regulation of protein apoptosis show a

Table 8: Distribution of successful drug target proteins in zones of BCPIN.

Zone	Number of proteins	Successful drug target	Total
1	1	1(0.95%)	8
2	105	3(2.43%)	
3	123	3(2.70%)	
4	111	1(3.57%)	
5	28	0(0%)	

zonal predominance central location. In fact, the central location zone is good for drug targets.

Table 9: Summary of the distribution of MAPK signalling cascade, positive regulation and negative regulation of signalling, apoptosis positive regulation and negative regulation of protein in zones of BCPIN.

Zone	MAPK signalling	Positive signalling	Negative signalling	Apoptosis	Positive apoptosis	Negative apoptosis
1	3(2.85%)	4 (20.05%)	5 (4.76%)	0 (0%)	0 (0%)	0 (0%)
2	3 (2.43%)	4 (8.29%)	4 (3.25%)	1 (0.8%)	0 (0%)	1 (0.81%)
3	4 (3.60%)	4 (2.33%)	6 (5.40%)	0 (0%)	0 (0%)	0 (0%)
4	0 (0%)	0 (0.80%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
5	0 (0%)	1 (4.76%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

3.4. The central zones of LCPIN exhibit functional specialization

A pathway enrichment screening was conducted to assess the functional importance of the identified topological patterns. The results showed that the zones near to the centre contain important functions for LCPIN. Moreover, these zones exhibited functional specialization indicated by a high concentration of proteins near the centre (see Table ??). Zone 1 regulates Gene Expression, Signaling, Immune system function, G1-G1/S Mitosis, G2-G2/M Mitosis, Ribosome, Eukaryotic Translation Initiation, Regulation of mRNA stability by proteins that bind AU-rich elements, a marked accumulation of relevant proteins was shown. This zone is thought to be a central hub, assumes a crucial involvement in various biological processes, and has significant importance in both molecular and cellular biology. It governs fundamental mechanisms that regulate gene expression, cellular signaling, immune system function, cell cycle progression, protein synthesis, and mRNA stability. Zone 2 was also shown to be enriched for these pathways, albeit to a lesser extent. In contrast, zones 3, 4 and 5 had lower enrichment levels compared to zones closer to the centre. Furthermore, important process in cellular respiration, such as oxidative phosphorylation, is governed by zones 4 and 5. These zones have exclusive control over specific pathways that are not observable in any other areas. In fact, some pathways show a decrease in the proportion of proteins in them as we move from the centre towards the periphery and vice versa.

Table 10: Summary of the distribution of MAPK signalling cascade, positive regulation and negative regulation of signalling, apoptosis positive regulation and negative regulation of protein in zones of BCPIN.

Pathways with significant enrichment	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5
Gene Expression	87.61%	28.68%	18.91%	-	-
Metabolism of amino acids and derivatives	81.90%	-	-	-	-
Metabolism of proteins	88.57%	20.38%	26.12%	-	-
Signal Transduction	21.90%	28.45%	17.11%	-	-
Eukaryotic Translation Initiation	73.33%	-	-	-	-
Ribosome	72.38%	-	-	-	-
Mitotic G1-G1/S phases	12.38%	-	-	-	-
Mitotic G2-G2/M phases	12.38%	5.82%	-	-	-
Regulation of mRNA stability by proteins that bind AU-rich elements	13.33%	5.69%	-	-	-
Immune System	20%	18.69%	33.33%	-	-
Metabolism	82.85%	-	-	29.62%	80.95%
Respiratory electron transport	-	-	-	14.28%	76.19%
Oxidative phosphorylation	-	-	-	14.28%	57.14%

4. Discussion

In this article, we modeled LCPIN as a metric space and created a network structure, classifying proteins into different zones based on their graph-theoretic distance to the central protein. Our network comprises 409 proteins and 4473 interactions. Analysis revealed RPS27A as the central protein. This protein plays a critical role in various cellular processes, including ribosome formation, translation regulation, cell proliferation, apoptosis regulation, cancer development, and cellular stress response [38] [25] [48] [36] [43]. Additionally, RPS27A interacts with proteins involved in MAPK signaling, cell cycle regulation, and signal transduction. This suggests its potential involvement in multiple aspects of cancer development and progression. Notably, MAPK signaling pathways are frequently dysregulated in cancer, impacting cell growth, differentiation, and survival. RPS27A's involvement in signal modulation suggests its potential role in regulating various signaling pathways that influence cell proliferation, apoptosis, and other cellular processes. Furthermore, its association with cell cycle proteins implies its participation in cell division and proliferation regulation. Dysregulation of cell cycle proteins can lead to uncontrolled cell growth and tumor formation. Overall, our findings suggest that RPS27A is a multifaceted protein with significant roles in cancer. Its overexpression in various cancers and involvement in crucial cellular processes highlight its potential as a promising candidate for therapeutic interventions. Additionally, its association with drug resistance and poor prognosis underscores its importance in clinical outcomes. Further research is necessary to elucidate the mechanisms underlying the role of RPS27A in cancer and explore its therapeutic potential.

Approximately 87% of proteins reside in zones 1 to 3. The average connectivity degree is 60%, 12.5%, and 4.5% for zones 1, 2, and 3, respectively. This highlights the dense core and sparse periphery of LCPIN, with proteins closer to the center exhibiting higher connectivity. Nodes farther from the center tend to have lower connectivity degrees, with some even being isolated (degree 1). Zones 1 to 3, closer to the network core, exhibit a higher concentration of proteins essential for various cellular functions, including signaling, cell cycle regulation, and apoptosis. These protein types are crucial in cancer biology and

hold promise as potential drug targets. For example, U2AF1 and SF3B1 are oncogenes involved in pre-mRNA splicing, with mutations linked to various cancers [6] [2]. Similarly, H3F3A, another oncogene, plays a vital role in chromatin remodeling, gene expression, and cellular differentiation. MCL1, an anti-apoptotic protein, regulates cell survival and inhibits apoptosis. Its overexpression is associated with tumor progression and resistance to chemotherapy in various cancers [23]. The SFN protein participates in diverse cellular processes, including cell cycle regulation, signal transduction, and apoptosis, and has been implicated in various cancers [14]. RBBP7, a crucial component in chromatin remodeling factors, plays an essential role in cell cycle regulation, DNA repair, and transcriptional control. Its involvement in various cellular processes is linked to tumorigenesis and cancer progression [39] [42]. PTEN, a tumor suppressor protein, plays a critical role in cell survival, proliferation, and migration. Its function is often lost through mutation or deletion, which is frequently observed in various types of cancer [18]. Notably, zones 1 to 3 display a high concentration of documented successful drug targets in the literature. This finding strongly suggests the presence of abundant protein depots within these zones with potential therapeutic applications.

Pathway enrichment analysis reveals distinct functional specializations across different zones. For example, Zone 1 exhibits a higher concentration of proteins associated with essential pathways, including the eukaryotic translation initiation process. This process regulates gene expression, ensuring accurate protein synthesis, controlling protein production, enabling cellular adaptation and response, influencing cell differentiation and development, and has been implicated in various disease mechanisms [15]. Additionally, Zone 1 plays a key role in amino acid metabolism and their derivatives, which are essential for protein synthesis, energy production, biosynthesis of essential molecules, detoxification, and maintaining nitrogen balance – all crucial for overall health and cellular function [40].

Ribosomes, also found in Zone 1, are essential for protein synthesis, a fundamental process in all living organisms, supporting cellular activities, growth, development, and survival [26]. Furthermore, the G1 and G1/S phases of the cell cycle, present in this zone, are critical for cell growth, DNA repair, replication initiation, checkpoint control, cellular differentiation, and development. These phases ensure accurate transmission of genetic material and proper functioning of cells within an organism [28].

In contrast, Zone 2 is characterized by the regulation of mRNA stability by proteins that bind AU-rich elements. This process is related to post-translational protein modification, which can significantly influence cancer development and progression [13]. Finally, Zones 4 and 5 are primarily associated with oxidative phosphorylation, a vital function for cancer. This process is necessary for ATP production, providing energy for the growth and proliferation of cancer cells. Notably, dysfunctional oxidative phosphorylation has been observed across various cancer forms, often associated with aggressive and treatment-resistant tumors [35]. Additionally, the respiratory electron transport chain, also present in these zones, plays a crucial role in ATP production, regenerating electron carriers, managing reactive oxygen species, and maintaining redox balance and regulation. Without this critical process, cells would be unable to efficiently produce energy and carry out essential metabolic functions [35].

In summary, the analysis of pathway enrichment reveals functional specialization within different zones. Zone 1 displays a higher concentration of proteins associated with diverse pathways, such as the eukaryotic translation initiation process, amino acid metabolism, and the cell cycle. Maintaining proper cellular function relies heavily on these processes. Zone 2, on the other hand, is characterized by mRNA stability regulation and post-translational protein modification, potentially impacting cancer development and progression. Finally, Zones 4 and 5 highlight the critical role of oxidative phosphorylation and the respiratory electron transport chain in cancer, both of which are linked to aggressive and treatment-resistant phenotypes. These findings strongly support the potential therapeutic significance of the protein depots within these zones.

5. Conclusions

This study utilizes a metric space modeling approach to construct and explore the topological properties of the lung cancer protein interaction network. Our analysis uncovers RPS27A as a central hub within the network, implicated in diverse cellular processes like ribosome biogenesis, translational regulation, cell proliferation, apoptosis control, and oncogenesis. This central position suggests potential pleiotropic roles for RPS27A in cancer progression, particularly in MAPK signaling, signal transduction, and cell cycle control. The observed spatial distribution of essential, signaling, cell cycle, and apoptosis-related proteins within central network zones highlights their potential significance in cancer biology and as potential druggable targets. Notably, the presence of established drug targets within these zones further bolsters the therapeutic potential of LCPIN proteins. However, further investigation is necessary to elucidate the precise mechanisms underlying their involvement in cancer and translate this knowledge into effective therapeutic strategies. Additionally, pathway enrichment analysis reveals distinct functional specializations within different network zones, providing valuable insights into the functional landscapes contributing to cellular dysfunction and disease progression. These findings establish a robust framework for future studies and pave the way for the development of novel therapeutic interventions targeting specific network zones and pathways. In conclusion, this study unveils the multifaceted roles of central LCPIN proteins and underscores the network's potential as a valuable resource for identifying novel therapeutic targets in lung cancer.

References

- [1] Albert-László Barabási, Natali Gulbahce, and Joseph Loscalzo. Network medicine: a network-based approach to human disease. *Nature reviews genetics*, 12(1):56–68, 2011.
- [2] Sebastian Bender, Yujie Tang, Anders M Lindroth, Volker Hovestadt, David TW Jones, Marcel Kool, Marc Zapatka, Paul A Northcott, Dominik Sturm, Wei Wang, et al. Reduced h3k27me3 and dna hypomethylation are major drivers of gene ex-

- pression in k27m mutant pediatric high-grade gliomas. *Cancer cell*, 24(5):660–672, 2013.
- [3] Peter Blume-Jensen and Tony Hunter. Oncogenic kinase signalling. *Nature*, 411(6835):355–365, 2001.
- [4] Victor Bryant. *Metric spaces: iteration and application*. Cambridge University Press, 1985.
- [5] Carol J Bult, Janan T Eppig, Judith A Blake, James A Kadin, Joel E Richardson, and Mouse Genome Database Group. The mouse genome database: genotypes, phenotypes, and models of human disease. *Nucleic acids research*, 41(D1):D885–D891, 2012.
- [6] Mario Cazzola, Marianna Rossi, and Luca Malcovati. Biologic and clinical significance of somatic mutations of sf3b1 in myeloid and lymphoid neoplasms. *Blood, The Journal of the American Society of Hematology*, 121(2):260–269, 2013.
- [7] Emad Fadhal. A comprehensive analysis of the pi3k/akt pathway: Unveiling key proteins and therapeutic targets for cancer treatment. *Cancer Informatics*, 22:11769351231194273, 2023.
- [8] Emad Fadhal. Unveiling crucial proteins in zone 2 of the human protein interaction network: Implications for cancer research and therapeutic targeting. *The Open Bioinformatics Journal*, 16(1), 2023.
- [9] Emad Fadhal, Junaid Gamiieldien, and Eric C Mwambene. Protein interaction networks as metric spaces: a novel perspective on distribution of hubs. *BMC systems biology*, 8(1):1–11, 2014.
- [10] Emad Fadhal, Eric C Mwambene, and Junaid Gamiieldien. Modelling human protein interaction networks as metric spaces has potential in disease research and drug target discovery. *BMC Systems Biology*, 8:1–12, 2014.
- [11] E Foulstone, S Prince, O Zaccheo, JL Burns, J Harper, C Jacobs, D Church, and AB Hassan. Insulin-like growth factor ligands, receptors, and binding proteins in cancer. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, 205(2):145–153, 2005.
- [12] Alvaro Gutierrez, Hannah Demond, Priscilla Brebi, and Carmen Gloria Ili. Novel methylation biomarkers for colorectal cancer prognosis. *Biomolecules*, 11(11):1722, 2021.
- [13] Douglas Hanahan and Robert A Weinberg. Hallmarks of cancer: the next generation. *cell*, 144(5):646–674, 2011.

- [14] Heiko Hermeking, Christoph Lengauer, Kornelia Polyak, Tong-Chuan He, Lin Zhang, Sam Thiagalingam, Kenneth W Kinzler, and Bert Vogelstein. 14-3-3 σ is a p53-regulated inhibitor of g2/m progression. *Molecular cell*, 1(1):3–11, 1997.
- [15] Praveen Kumar Jaiswal, Sweaty Koul, Nallasivam Palanisamy, and Hari K Koul. Eukaryotic translation initiation factor 4 gamma 1 (eif4g1): a target for cancer therapeutic intervention? *Cancer Cell International*, 19(1):1–14, 2019.
- [16] Silvia Lapenna and Antonio Giordano. Cell cycle kinases as therapeutic targets for cancer. *Nature reviews Drug discovery*, 8(7):547–566, 2009.
- [17] Eva YHP Lee and William J Muller. Oncogenes and tumor suppressor genes. *Cold Spring Harbor perspectives in biology*, 2(10):a003236, 2010.
- [18] Jing Li, Clifford Yen, Danny Liaw, Katrina Podsypanina, Shikha Bose, Steven I Wang, Janusz Puc, Christa Miliarexis, Linda Rodgers, Richard McCombie, et al. Pten, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *science*, 275(5308):1943–1947, 1997.
- [19] Zhaopeng Li, Deyong Bu, Xiaobin Wang, Lin Zhu, Daoyan Lei, Fengling Tang, Xi-anhua Sun, Cheng Chen, Xiang Ji, and Song Bai. Chidamide and oxaliplatin synergistically inhibit colorectal cancer growth by regulating the rps27a-mdm2-p53 axis. *OncoTargets and Therapy*, pages 703–721, 2023.
- [20] Junxia Min, Alexander Zaslavsky, Giuseppe Fedele, Sara K McLaughlin, Elizabeth E Reczek, Thomas De Raedt, Isil Guney, David E Storchlic, Laura E MacConaill, Rameen Beroukhim, et al. An oncogene–tumor suppressor cascade drives metastatic prostate cancer by coordinately activating ras and nuclear factor- κ b. *Nature medicine*, 16(3):286–294, 2010.
- [21] Karl Münger. Disruption of oncogene/tumor suppressor networks during human carcinogenesis. *Cancer investigation*, 20(1):71–81, 2002.
- [22] Francesca Nazio, Matteo Bordi, Valentina Cianfanelli, Franco Locatelli, and Francesco Cecconi. Autophagy and cancer stem cells: molecular mechanisms and therapeutic applications. *Cell Death & Differentiation*, 26(4):690–702, 2019.
- [23] Rhonda M Perciavalle and Joseph T Opferman. Delving deeper: Mcl-1’s contributions to normal and cancer biology. *Trends in cell biology*, 23(1):22–29, 2013.
- [24] Laura Francesca Pisani, Isabella Teani, Maurizio Vecchi, and Luca Pastorelli. Interleukin-33: Friend or foe in gastrointestinal tract cancers? *Cells*, 12(11):1481, 2023.
- [25] Giuseppina Raspaglio, Flavia Filippetti, Silvia Prislei, Roberta Penci, Ilaria De Maria, Lucia Cicchillitti, Simona Mozzetti, Giovanni Scambia, and Cristiano Ferlini. Hypoxia induces class iii beta-tubulin gene expression by hif-1 α binding to its 3’flanking region. *Gene*, 409(1-2):100–108, 2008.

- [26] Marina V Rodnina and Wolfgang Wintermeyer. Fidelity of aminoacyl-trna selection on the ribosome: kinetic and structural mechanisms. *Annual review of biochemistry*, 70(1):415–435, 2001.
- [27] Thomas Rolland, Murat Taşan, Benoit Charloteaux, Samuel J Pevzner, Quan Zhong, Nidhi Sahni, Song Yi, Irma Lemmens, Celia Fontanillo, Roberto Mosca, et al. A proteome-scale map of the human interactome network. *Cell*, 159(5):1212–1226, 2014.
- [28] Concepción Sánchez-Martínez, María José Lallena, Sonia Gutiérrez Sanfeliciano, and Alfonso de Dios. Cyclin dependent kinase (cdk) inhibitors as anticancer drugs: Recent advances (2015–2019). *Bioorganic & medicinal chemistry letters*, 29(20):126637, 2019.
- [29] Judith S Sebolt-Leopold and Roman Herrera. Targeting the mitogen-activated protein kinase cascade to treat cancer. *Nature reviews cancer*, 4(12):937–947, 2004.
- [30] Rebecca L Siegel, Kimberly D Miller, Hannah E Fuchs, and Ahmedin Jemal. Cancer statistics, 2022. *CA: a cancer journal for clinicians*, 72(1):7–33, 2022.
- [31] Asad Ullah, Zohra Bano, and Shahid Zaman. Computational aspects of two important biochemical networks with respect to some novel molecular descriptors. *Journal of Biomolecular Structure and Dynamics*, 42(2):791–805, 2024.
- [32] Asad Ullah, Safina Jabeen, Shahid Zaman, Anila Hamraz, and Summeira Meherban. Predictive potential of k-banhatti and zagreb type molecular descriptors in structure–property relationship analysis of some novel drug molecules. *Journal of the Chinese Chemical Society*, 2024.
- [33] Asad Ullah, Shahid Zaman, Anila Hamraz, and Muniba Muzammal. On the construction of some bioconjugate networks and their structural modeling via irregularity topological indices. *The European Physical Journal E*, 46(8):72, 2023.
- [34] Bert Vogelstein, Nickolas Papadopoulos, Victor E Velculescu, Shibin Zhou, Luis A Diaz Jr, and Kenneth W Kinzler. Cancer genome landscapes. *science*, 339(6127):1546–1558, 2013.
- [35] Douglas C Wallace, Weiwei Fan, and Vincent Procaccio. Mitochondrial energetics and therapeutics. *Annual Review of Pathology: Mechanisms of Disease*, 5:297–348, 2010.
- [36] Jing Wan, Juan Lv, Chun Wang, and Li Zhang. Rps27 selectively regulates the expression and alternative splicing of inflammatory and immune response genes in thyroid cancer cells. *Advances in Clinical and Experimental Medicine*, 31(8):889–901, 2022.
- [37] Houcai Wang, Bingqian Xie, Yuanyuan Kong, Yi Tao, Guang Yang, Minjie Gao, Hongwei Xu, Fenghuang Zhan, Jumei Shi, Yiwen Zhang, et al. Overexpression of rps27a contributes to enhanced chemoresistance of cml cells to imatinib by the trans-activated stat3. *Oncotarget*, 7(14):18638, 2016.

- [38] Limei Wang, Jin Zhang, Cunwei Wang, Yuping Qi, Mi Du, Wenhua Liu, Chengzhe Yang, and Pishan Yang. Low concentrations of $\text{tnf-}\alpha$ promote osteogenic differentiation via activation of the ephrinb2-ephb4 signalling pathway. *Cell proliferation*, 50(1):e12311, 2017.
- [39] Renfeng Wang, Zhiliang Huang, Zhenyang Lin, Borong Chen, Xuguang Pang, Chaoxiang Du, and Hong Fan. Hypoxia-induced rbbp7 promotes esophagus cancer progression by inducing cdk4 expression: Rbbp7 promotes esophagus cancer progression. *Acta Biochimica et Biophysica Sinica*, 54(2):179, 2022.
- [40] Kathryn E Wellen and Craig B Thompson. A two-way street: reciprocal regulation of metabolism and signalling. *Nature reviews Molecular cell biology*, 13(4):270–276, 2012.
- [41] Guanming Wu, Xin Feng, and Lincoln Stein. A human functional protein interaction network and its application to cancer data analysis. *Genome biology*, 11:1–23, 2010.
- [42] Z-F Xie, H-T Li, S-H Xie, and M Ma. Circular rna hsa_circ.0006168 contributes to cell proliferation, migration and invasion in esophageal cancer by regulating mir-384/rbbp7 axis via activation of s6k/s6 pathway. *European Review for Medical & Pharmacological Sciences*, 24(1), 2020.
- [43] Kenichi Yoshida, Masashi Sanada, Yuichi Shiraishi, Daniel Nowak, Yasunobu Nagata, Ryo Yamamoto, Yusuke Sato, Aiko Sato-Otsubo, Ayana Kon, Masao Nagasaki, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, 478(7367):64–69, 2011.
- [44] Shahid Zaman, Mehwish Jalani, Asad Ullah, Mubashir Ali, and Tayyba Shahzadi. On the topological descriptors and structural analysis of cerium oxide nanostructures. *Chemical Papers*, 77(5):2917–2922, 2023.
- [45] Shahid Zaman, Hafiza Sana Arbab Yaqoob, Asad Ullah, and Mariam Sheikh. Qspr analysis of some novel drugs used in blood cancer treatment via degree based topological indices and regression models. *Polycyclic Aromatic Compounds*, pages 1–17, 2023.
- [46] Wei Zhou, Jiarui Wu, Jingyuan Zhang, Xinkui Liu, Siyu Guo, ShanShan Jia, Xiaomeng Zhang, Yingli Zhu, and Miaomiao Wang. Integrated bioinformatics analysis to decipher molecular mechanism of compound kushen injection for esophageal cancer by combining wgcna with network pharmacology. *Scientific reports*, 10(1):12745, 2020.
- [47] Weiyu Zhou, Yujing Wang, Hongmei Gao, Ying Jia, Yuanxin Xu, Xiaojing Wan, Zhiying Zhang, Haiqiao Yu, and Shuang Yan. Identification of key genes involved in pancreatic ductal adenocarcinoma with diabetes mellitus based on gene expression profiling analysis. *Pathology and Oncology Research*, 27:604730, 2021.

- [48] Xiang Zhou, Wen-Juan Liao, Jun-Ming Liao, Peng Liao, and Hua Lu. Ribosomal proteins: functions beyond the ribosome. *Journal of molecular cell biology*, 7(2):92–104, 2015.