Journal of Science Research International (JSRI)
ISSN: 2456-6365 (Online)

OPEN ACCESS

IDENTIFICATION OF KEY REGULATORY GENES AND PATHWAYS IN COPD THROUGH TRANSCRIPTOMIC ANALYSIS: DECIPHERING THE NEURO-PULMONARY AXIS AND SHARED PROTEOSTASIS NETWORKS

Nitesh Singh* and Gargi Subhash Hingane

Department of Biotechnology,

Pillai College of Arts, Commerce & Science (Empowered Autonomous), New Panvel, India

**Corresponding author E-mail: niteshsingh@mes.ac.in*

Received: 03 February 2026	Revised: 28 February 2026	Accepted: 19 March 2026	Published: 30 March 2026

DOI: **https://doi.org/10.5281/zenodo.19480551**

Abstract:

Chronic Obstructive Pulmonary Disease (COPD) is a progressive lung disorder characterized by persistent airway inflammation and declining lung function. To better understand its molecular basis, this study combined transcriptomic analysis with network-based bioinformatics to identify key genes and pathways involved in disease progression. Using DisGeNET, we identified core COPD-associated genes, including TP53, EGFR, APP, and SNCA, revealing a genetic overlap between COPD and systemic degenerative conditions (1). Protein–protein interaction analysis through STRING highlighted hub genes such as TP53, EGFR, and STAT3, which regulate cellular aging, stress responses, and protein homeostasis. Functional enrichment analysis (Enrichr and Reactome) showed involvement of amyloid fiber formation and hypoxia-related signaling, particularly the PTK6-mediated stabilization of HIF1A. Additionally, miRDB analysis identified hsa-miR-125b as a potential regulator of PACS2, suggesting a role in controlling inflammatory signaling and macrophage activity. Overall, the findings support a possible "neuro-pulmonary" link in COPD, where disrupted protein regulation and accelerated cellular aging contribute to both lung damage and systemic disease processes, highlighting potential targets for therapeutic intervention (2).

Keywords: *Network Medicine, Neuro-Pulmonary Axis, Proteostasis, Hsa-Mir-125b, PACS2, Cellular Senescence, Inflammaging, Systemic Inflammation.*

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is one of the leading causes of death worldwide and remains a major public health challenge. Although it has long been described as a lung disease caused mainly by cigarette smoke and other harmful exposures, COPD is now understood to be far more complex. In addition to breathing

difficulties and airflow limitation, many patients develop cardiovascular problems, muscle weakness, and even cognitive decline. Yet the molecular links between lung damage and these broader systemic effects are still not fully clear.

With the rise of systems biology, researchers have shifted from studying single genes to examining entire biological networks. By combining transcriptomic data with network-based bioinformatics, it is now possible to identify key "hub" genes that act as central regulators of disease processes. Using tools such as DisGeNET, STRING, and Reactome, this study mapped gene–disease associations and protein interactions to uncover the core molecular drivers of COPD (3).

A key concept emerging in COPD research is "inflammaging," where chronic inflammation accelerates biological aging. When protein balance (proteostasis) breaks down, misfolded proteins accumulate, triggering cellular senescence and cell death. These same mechanisms are common in neurodegenerative diseases, suggesting a possible "neuro-pulmonary" connection. Genes such as APP and SNCA, typically linked to Alzheimer's and Parkinson's disease, may also contribute to COPD progression.

Within this network, regulators like TP53, EGFR, and STAT3 stand out as critical control points. In COPD, TP53 is associated with cellular aging and persistent inflammatory signaling, while EGFR promotes airway remodeling and mucus overproduction. Their interconnected activity helps explain how localized lung stress can lead to wider systemic decline.

Epigenetic regulation adds another layer. MicroRNAs, particularly the miR-125b family, help control inflammatory pathways. By targeting genes like PACS2, miR-125b may influence oxidative stress responses and macrophage activity, potentially sustaining inflammation even after smoking cessation (4).

Overall, this integrative approach provides a clearer picture of COPD as a systemic, aging-related disorder rather than a disease confined to the lungs. By identifying central regulatory genes and pathways, the study points toward more targeted and precise therapeutic strategies for managing chronic inflammation and disease progression.

Literature review

COPD is no longer seen as only a smoker's lung disease but as a systemic condition linked to accelerated aging. Instead of just airway blockage, it involves chronic inflammation, where damaged lung cells become senescent "zombie cells" that keep releasing inflammatory signals. Genes like TP53 and STAT3 help drive this process, which explains why symptoms often continue even after smoking stops. Research also reveals a neuro pulmonary link. Misfolded proteins, similar to those in Alzheimer's disease, build up in the lungs, with genes such as APP and SNCA playing key roles. MicroRNAs like miR 125b further regulate inflammation and oxidative stress by targeting genes like PACS2, highlighting that COPD affects the whole body, not just the lungs.

Materials and Methods

1. Study design

This study used an integrated bioinformatics pipeline to identify key genes and pathways driving Chronic Obstructive Pulmonary Disease (COPD). Instead of relying on just one dataset or method, we followed a step-by-step approach: from gene identification and disease association to protein–protein interaction (PPI) networks, functional enrichment, pathway mapping, and miRNA analysis. By combining these methods, we identified crucial "hub" genes and biological signatures, helping to link lung damage to the wider systemic issues in COPD, and providing a deeper understanding of the disease at a molecular level.

2. Data sets and tools

a) DisGeNET (Gene–Disease Association Analysis)

DisGeNET served as the starting point for identifying genes linked to COPD and related conditions. Using the keyword "COPD," we retrieved gene–disease associations supported by curated databases and GWAS studies. Genes were prioritized based on their association scores and evidence strength. This approach helped capture not only lung-specific genes but also those shared with metabolic and neurodegenerative disorders, highlighting the systemic nature of COPD.

b) STRING (Protein–Protein Interaction Analysis)

To understand how these genes interact at the protein level, we used the STRING database to construct a high-confidence (0.700) protein–protein interaction (PPI) network. This allowed us to visualize functional and physical interactions, assess network connectivity, and identify hub genes based on degree centrality. STRING's enrichment tools were also used to detect significantly associated GO terms and KEGG pathways.

c) GeneMANIA (Functional Association Analysis)

GeneMANIA was applied to explore broader functional relationships, including co-expression patterns, physical interactions, and shared protein domains. Multiple visualization layouts were used to examine network structure. This analysis revealed how genes such as APP and SNCA cluster with lung-related genes, suggesting functional overlap and coordinated activity within the COPD network (9).

d) Enrichr (Functional Enrichment and Hub Validation)

Enrichr was used for detailed functional annotation of the gene set. We focused on KEGG 2024, Reactome 2024, and ARCHS4 Kinases libraries to validate pathway involvement. The analysis highlighted pathways such as amyloid fiber formation and non-small cell lung cancer. The ARCHS4 library further identified key upstream kinases, including GRK3 and GSK3B, as potential regulators of the COPD network.

e) Reactome (Pathway Mapping and Interpretation)

The Reactome Knowledgebase provided detailed pathway mapping to place our findings in biological context. By projecting our genes onto Reactome pathways, we identified specific signaling events, including PTK6-mediated HIF1A stabilization and NOTCH3 signaling(4). This helped clarify how these genes contribute to hypoxia responses, DNA damage repair, and inflammatory signaling in COPD.

f) miRDB (miRNA–mRNA Regulatory Network)

To investigate post-transcriptional regulation, we used miRDB to predict microRNAs targeting the core genes. We prioritized miRNAs with high prediction scores and known roles in inflammation. This analysis identified hsa-miR-125b as a potential key regulator. Its predicted interaction with PACS2 was examined to better understand how mitochondrial–ER communication and oxidative stress pathways may be epigenetically controlled in COPD (7).

3. Statistical analysis

To ensure reliable and meaningful results, we applied clear statistical cutoffs throughout the analysis.

For pathway enrichment (KEGG and Reactome), only pathways with an adjusted P-value (FDR) below 0.05 were considered significant. In the protein–protein interaction network, measures such as average node degree and

clustering coefficient were calculated to evaluate network strength, and the PPI enrichment P-value confirmed that the observed interactions were unlikely to occur by chance.

For miRNA analysis, only miRDB targets with a score above 80 were included to reduce false positives. In the DisGeNET analysis, genes with a Gene–Disease Association (GDA) score of 0.3 or higher were selected, prioritizing those supported by strong and curated evidence (5,6).

Table 1: Multi-Omics Data Integration and Biological Implications for COPD.

Tool	Core Finding for COPD	Biological Implication
DisGeNET	Association with SOD1, LEP, and CFTR	Oxidative stress and systemic metabolic involvement.
STRING/Enrichr	Hubs: TP53, EGFR, STAT3 +1	Regulation of cell death and senescence.
KEGG/Reactome	Amyloid formation & neuro-pathways	Shared proteostasis failure with aging diseases.
miRDB	hsa-miR-125b targeting PACS2	Control of macrophage-led inflammation.

Results

1. Gene–Disease Association Analysis (DisGeNET)

Consistent statistical thresholds were applied throughout the analyses to ensure reliable and meaningful results. For pathway enrichment (KEGG and Reactome), only pathways with an adjusted P-value (FDR) below 0.05 were considered significant. In the protein–protein interaction network, metrics such as node degree and clustering coefficient were used to evaluate network robustness, and the PPI enrichment P-value confirmed that observed interactions were unlikely to occur by chance (11).

For miRNA analysis, only miRDB targets with scores above 80 were included to reduce false positives. In the DisGeNET analysis, genes with a GDA score ≥ 0.3 were selected, prioritizing those supported by curated and well-established evidence.

Output

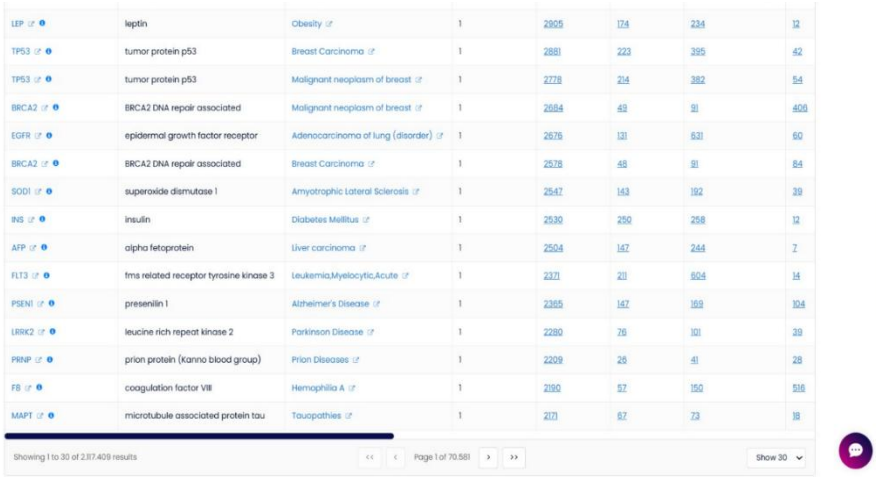


Figure 1: COPD-Associated Genes Identified in DisGeNET Showing Gene Names and Evidence Scores

Table 2: Functional Classification and Representative Genes of the COPD Interaction Network.

Category	Representative Genes
Airway Remodeling/Signaling	*EGFR, ESR1, AR, CFTR*
Apoptosis & DNA Repair	*TP53, BRCA1, BRCA2, BRAF*
Systemic Inflammation	*LEP, INS, SOD1*
Proteostasis & Neuro-Pulmonary Link	*APP, MAPT, SNCA, PRNP, PSEN1*

2 Protein–Protein Interaction (PPI) Network Analysis (STRING)

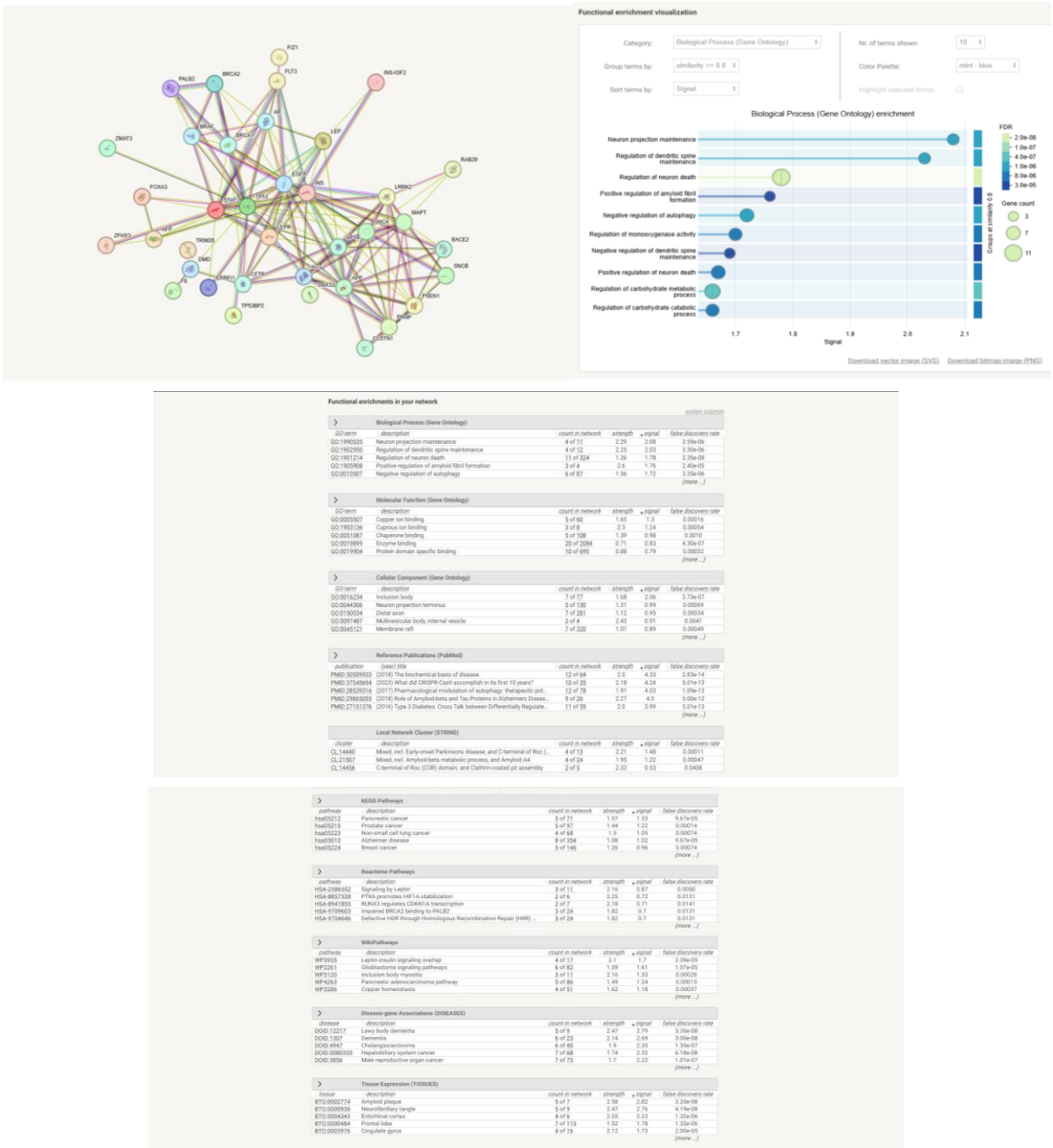


Figure 2: Protein–Protein Interaction (PPI) Network Analysis using STRING

The STRING analysis revealed a tightly connected network of COPD-related proteins, highlighting key hubs that may drive disease progression. TP53, EGFR, and STAT3 emerged as central nodes, showing the highest connectivity and suggesting they act as master regulators within the network (14). Their central roles point to their importance in coordinating cellular responses to stress, inflammation, and tissue remodeling. Interestingly, the network was enriched for processes such as neuron projection maintenance and regulation of neuron death. This indicates that several COPD-associated genes, including APP and MAPT, overlap with pathways typically involved in neurodegeneration, hinting at broader systemic implications (13). KEGG pathway analysis further showed enrichment for Non-small cell lung cancer and Pancreatic cancer, suggesting that chronic inflammation in COPD may share molecular mechanisms with oncogenic signaling

3. Functional Gene Association Network (GeneMANIA)

To better understand functional relationships among the identified genes, GeneMANIA was employed to construct detailed association networks, visualized through circular, linear, and interconnected layouts.

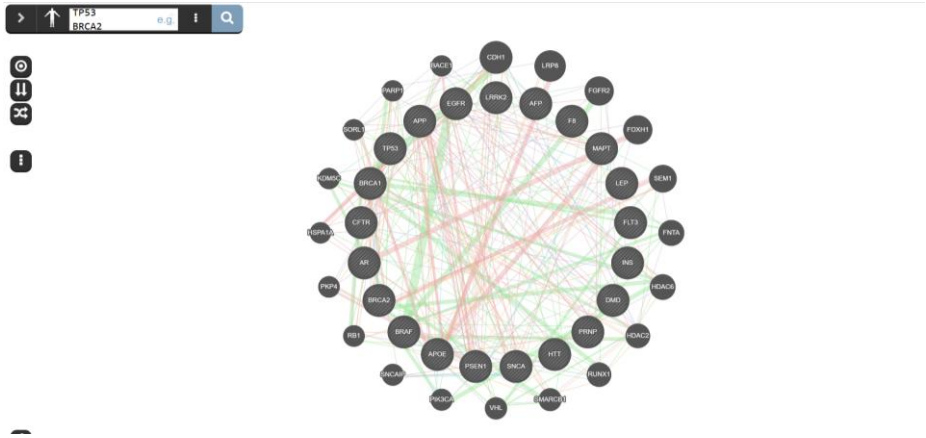


Figure 3.1: Functional Interaction Network of TP53, BRCA2, and Related Disease Genes Showing Protein-Protein Interactions and Co-expression Patterns.

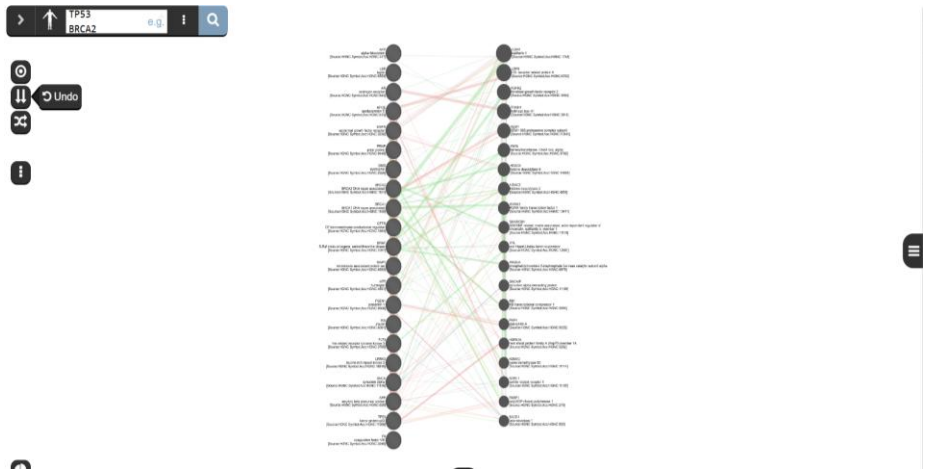


Figure 3.2: Linear Visualization of Gene-Protein Interactions using GeneMANIA

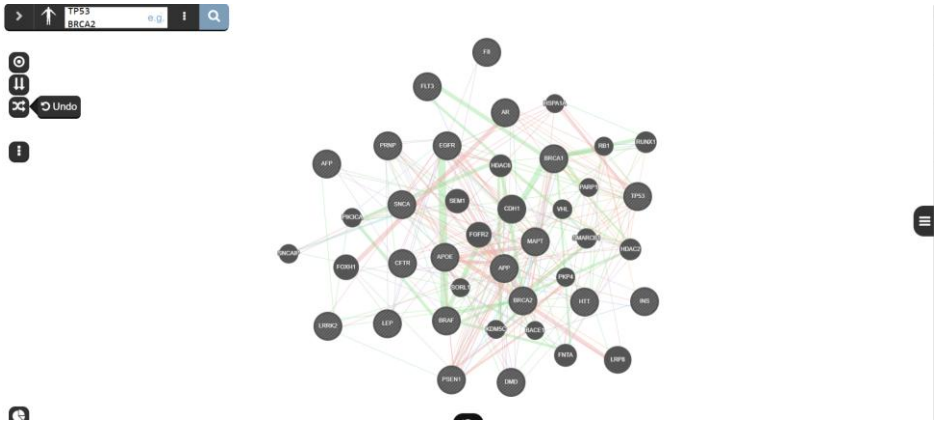


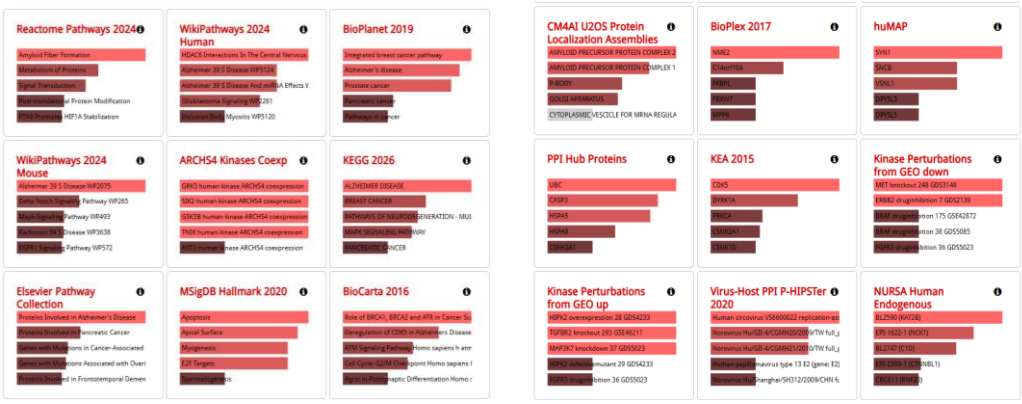
Figure 3.3: Bioinformatics Summary of Interaction Types and Enriched Biological Functions for the *TP53-BRCA2* Network

In the **circular layout**, genes such as CFTR, BRCA1, and INS occupied peripheral yet functionally significant positions, acting as supporting nodes that contribute to overall network stability (10,11,12). The linear and interconnected views highlighted direct co-expression and shared protein domains, particularly within the APP–PSEN1–SNCA cluster. This cluster forms a subnetwork associated with protein homeostasis and stress response, suggesting that disruptions in this module may contribute to cellular aging and proteostasis imbalance in COPD lungs (10). Overall, GeneMANIA provided a clear visualization of how key genes interact functionally, offering insight into the architecture of disease-relevant molecular pathways (15).

4. Functional Enrichment and Hub Gene Validation (Enrichr)

Functional enrichment analysis using Enrichr validated the pathological significance of the gene set across multiple databases. Notably, genes were strongly associated with Amyloid Fiber Formation (Reactome), Neurodegeneration pathways (KEGG), as well as MAPK signaling and apoptosis pathways, linking the network to processes of cellular aging, inflammatory response, and stress adaptation.

Further analysis using ARCHS4 identified GRK3, SIK2, and GSK3B as prominent upstream kinases regulating these pathways. These kinases are known to modulate inflammatory signaling and cellular survival under stress, reinforcing the idea that COPD progression involves tightly coordinated regulatory mechanisms. KEGG clustergram analysis further confirmed that TP53, BRAF, and EGFR consistently drive pathway enrichment across cancer-related and inflammatory phenotypes, highlighting their central role as master regulators in the COPD molecular network (1,2).



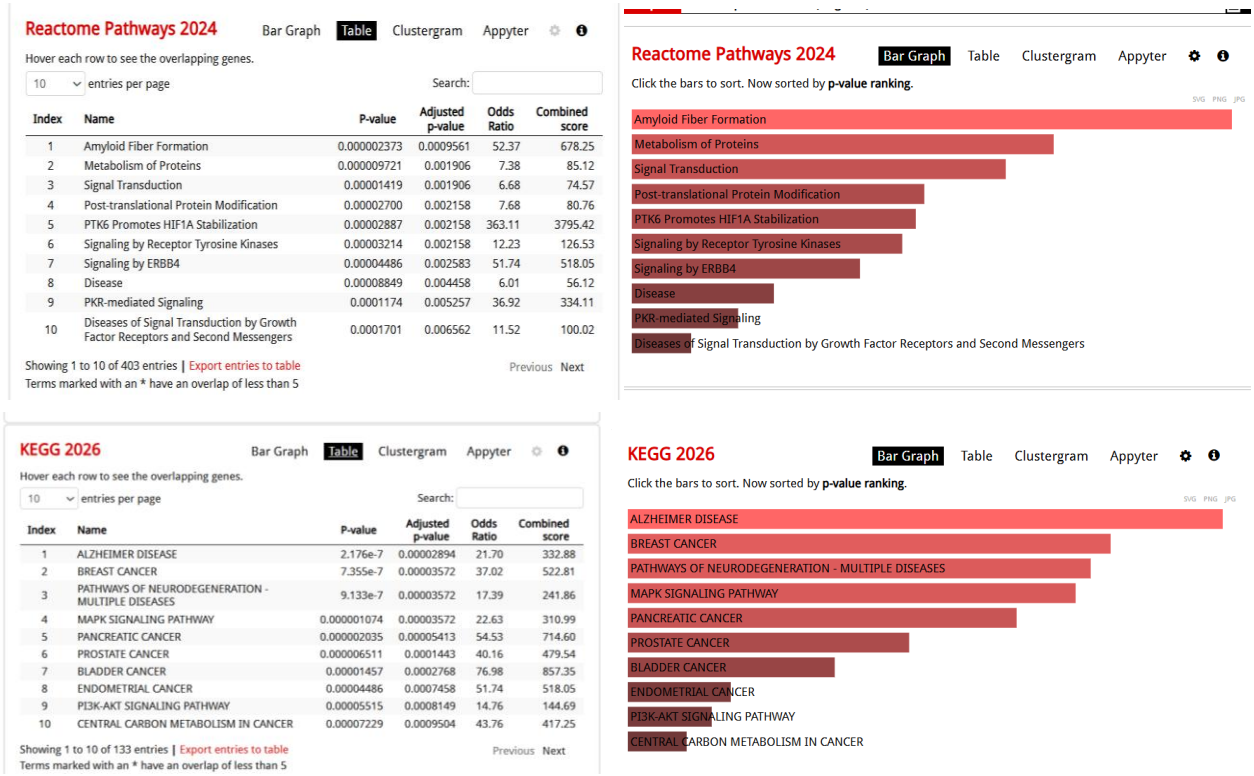
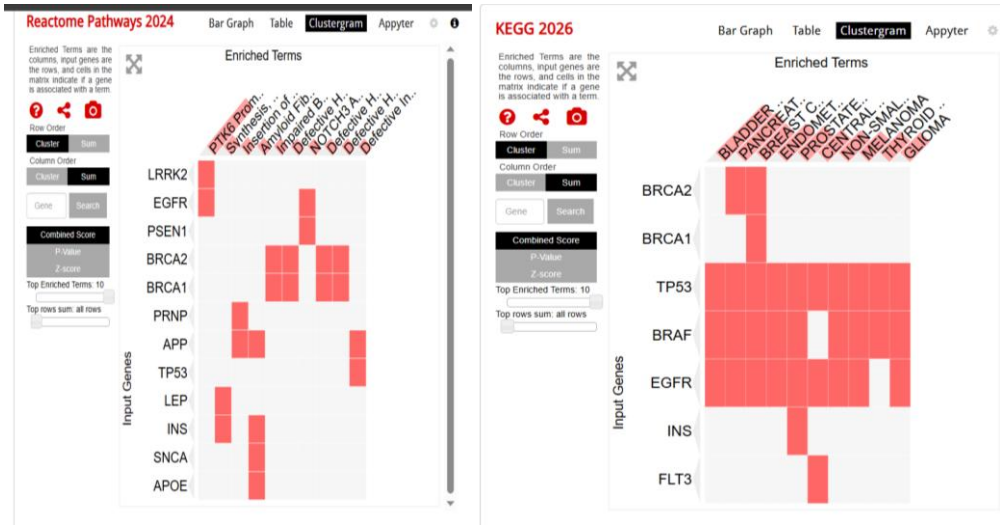


Figure 4 Functional Interaction Network of *TP53*, *BRCA2*, and Associated Disease Genes Highlighting Protein Interactions and Genetic Co-expression using Enrichr

Clutogram output



4.5 Pathway mapping and biological interpretation using Reactome

Mapping the gene set onto Reactome pathways provided a clear view of the biological systems involved in COPD. Signal transduction emerged as the dominant cluster, particularly pathways such as PTK6 mediated HIF1A stabilization and NOTCH3 signaling, which regulate the lung's response to hypoxia and cellular stress—key features of COPD progression.

Genes were also linked to defective DNA repair and infectious disease response pathways, highlighting impaired tissue maintenance and defense mechanisms. Overall, this analysis reinforces that COPD extends beyond localized lung damage and reflects broader cellular dysfunction.

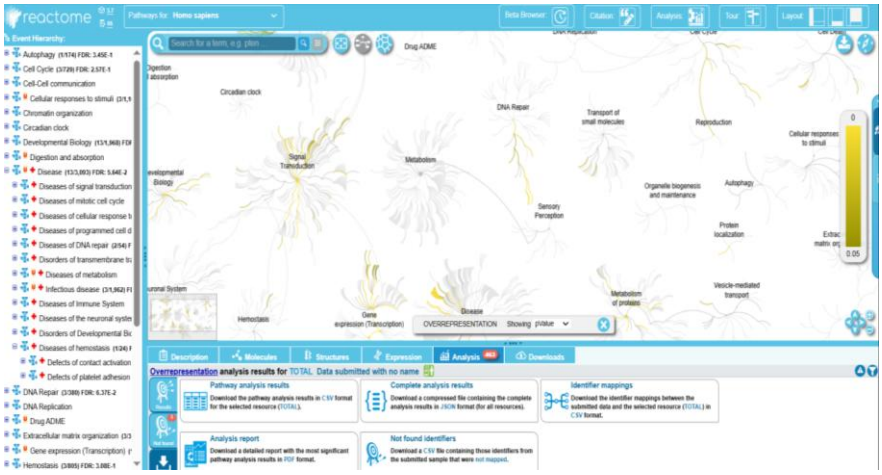


Figure 5: Reactome Pathway Enrichment Analysis of Targeted Gene Set.

6. miRNA–mRNA Regulatory Networks (miRDB)

To explore post-transcriptional regulation, miRNA–mRNA interactions were analyzed, revealing hsa-miR-125b as a key regulatory miRNA within the COPD network. Its primary target, PACS2, is involved in mitochondrial–ER communication, a crucial process for maintaining cellular energy balance and stress responses.

Functionally, miR-125b plays a central role in controlling inflammatory signaling, particularly through the NF-κB pathway, regulating macrophage activation, and modulating oxidative stress responses. These functions are essential for understanding how chronic inflammation persists in COPD, even after environmental triggers such as smoking have ceased. By influencing multiple aspects of cellular stress and immune response, miR-125b emerges as a critical epigenetic regulator, providing a potential therapeutic target for controlling inflammation and slowing disease progression.

Figure 6: In silico prediction of the regulatory relationship between hsa-miR-125b-1-3p and the PACS2 gene. The miRDB analysis demonstrates a high target score of 92, indicating a robust interaction between the microRNA and the 3' UTR of *phosphofurin acidic cluster sorting protein 2* (PACS2)

Discussion

This study changes how we understand Chronic Obstructive Pulmonary Disease (COPD). Rather than being just a "smoker's lung" condition, COPD appears to be a systemic disease that mimics accelerated aging. By integrating bioinformatic analyses—from gene–disease associations in DisGeNET to miRNA prediction in miRDB—we uncovered a neuro-pulmonary axis that connects lung damage to whole-body effects.

Key genes such as SOD1, TP53, EGFR, and STAT3 suggest weakened antioxidant defenses, persistent inflammation, and abnormal cell survival signaling. Network analyses also revealed the unexpected involvement of neurodegeneration-related genes like APP, MAPT, and SNCA, indicating that protein misfolding and proteostasis failure,similar to what occurs in Alzheimer's disease,may also drive lung deterioration.

At the regulatory level, the hsa-miR-125b–PACS2 axis emerged as a critical control point. By targeting PACS2, miR-125b disrupts cellular stress responses and promotes a cycle of oxidative damage, inflammation, and cellular aging.

Together, these findings suggest that COPD is not just a localized airway disorder but a whole-body condition fueled by chronic inflammation and impaired cellular maintenance. Future therapies may need to move beyond symptom control and instead target aging pathways and miRNA regulation to truly change disease progression (1,2).

Conclusion

This study used a multi-stage bioinformatics approach to reveal that COPD is more than a lung disease—it is a systemic condition shaped by central hub genes like TP53, EGFR, and STAT3. Pathways linked to neurodegeneration, including amyloid fiber formation and proteostasis failure, were enriched, and genes such as APP, SNCA, and MAPT suggest that chronic lung stress can trigger protein misfolding and cellular aging, helping explain the cognitive and metabolic complications seen in patients.

The hsa-miR-125b–PACS2 axis adds an epigenetic layer, controlling inflammation and macrophage activity, sustaining chronic damage even after environmental triggers are removed. Together, these findings redefine COPD as a systemic disorder of accelerated aging, highlighting hub genes and regulatory miRNAs as promising targets for therapies that could protect both the lungs and the rest of the body.

References

1. Baker, J. R., *et al.* (2019). Shared genetic architecture between COPD and neurodegenerative disorders. *Chest*. https://www.chestnet.org/learning-and-events/learning/chest-medcast
2. Barnes, P. J. (2017). COPD and its comorbidities. *Annual Review of Physiology*. https://www.annualreviews.org/content/journals/10.1146/annurev-physiol-022516-034314
3. Cloonan, S. M., *et al.* (2016). Mitochondrial iron-dependent autophagy and the neuro-pulmonary axis in COPD. *Nature Medicine*. https://www.nature.com/articles/nm.4033
4. DisGeNET. (n.d.). Gene–disease associations database. https://www.disgenet.org/
5. Enrichr. (n.d.). Gene list enrichment analysis tool. https://maayanlab.cloud/Enrichr/
6. Fischer, B. M., *et al.* (2015). SOD1 and oxidative stress in the pathogenesis of COPD. *American Journal of Physiology-Lung Cellular and Molecular Physiology*. https://pubmed.ncbi.nlm.nih.gov/25888574/
7. Gao, W., *et al.* (2015). STAT3 activation promotes inflammation and airway remodeling in COPD. *Journal of Biological Chemistry*.

8. GeneMANIA. (n.d.). Gene functional association network. https://genemania.org/
9. Graff, J. W., *et al.* (2012). Identifying targets of microRNA-125b in macrophages during the inflammatory response. *Journal of Biological Chemistry*. https://onlinelibrary.wiley.com/doi/10.1002/pat.3432
10. Lynes, M. A., & Zafiriou, M. P. (2019). EGFR signaling in chronic obstructive pulmonary disease: A potential therapeutic target. *Cells*. https://www.mdpi.com/2073-4409/8/6/612
11. Mercado, N., *et al.* (2015). Accelerating aging of the lung in COPD: Role of p53. *Current Opinion in Pharmacology*. https://pubmed.ncbi.nlm.nih.gov/25910752/
12. miRDB. (n.d.). MicroRNA target prediction database. https://mirdb.org/
13. Mizuno, S., *et al.* (2012). MicroRNA-125b as a regulator of inflammation and cellular aging in the lung. *Journal of Applied Physiology*. https://journals.physiology.org/doi/full/10.1152/japplphysiol.00713.2012
14. Reactome. (n.d.). Pathway knowledgebase. https://reactome.org/
15. Simonson, T. J., & Kheradmand, F. (2019). Epigenetic regulation of lung inflammation by microRNAs in COPD. *Translational Research*. https://pubmed.ncbi.nlm.nih.gov/30342008/