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

OPEN ACCESS

NETWORK-BASED SYSTEMS BIOLOGY ANALYSIS FOR TARGET AND PATHWAY IDENTIFICATION IN CYSTIC FIBROSIS

Suparna Deepak* and Archana Yadav

Department of Biotechnology,

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

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

Received: 02 February 2026	Revised: 28 February 2026	Accepted: 13 March 2026	Published: 30 March 2026

DOI: **https://doi.org/10.5281/zenodo.19472499**

Abstract:

Cystic fibrosis (CF) is a life-threatening autosomal recessive disorder due to mutations in the CFTR gene, resulting in abnormal chloride transport, airway inflammation, and progressive pulmonary damage. While CFTR dysfunction is the main molecular cause, disease severity is modulated by intricate gene interaction networks and inflammatory signaling pathways. In this study, a bioinformatics-based network analysis approach was used to uncover important target genes and pathways for CF. A list of CF-related seed genes was obtained from the DisGeNET database and supplemented by GeneMANIA to build a functional interaction network. Protein-protein interaction analysis was carried out using the STRING tool, and hub genes were selected according to network topology. Pathway enrichment analysis was performed using Reactome. A total of 16 high-confidence CF-associated genes were analyzed. and TNF, IL1B, and IL6 were found to be high-degree hub genes. Reactome enrichment analysis showed significant participation of Interleukin-1 signaling, cytokine-mediated immune response, neutrophil degranulation, and oxidative stress pathways. These results underscore the pivotal role of inflammatory and immune regulatory pathways in CF pathogenesis. This study offers a systems-level perspective of CF molecular pathogenesis and points to new therapeutic targets that could be used in addition to CFTR-modulator therapy.

Keywords: *Cystic Fibrosis, Network Biology, Hub Genes, Bioinformatics, Cytokine Signaling, Systems Biology.*

Introduction

Cystic fibrosis (CF) is an autosomal recessive disease due to mutations in the CF transmembrane conductance regulator (CFTR) gene, leading to abnormal chloride transport, dehydrated airway surface epithelia, and chronic pulmonary inflammation. Chronic neutrophilic pulmonary inflammation and progressive lung damage remain the major cause of morbidity and mortality in CF patients (1).

Although CF is a result of CFTR dysfunction, it has been increasingly acknowledged that the progression of CF can be attributed to the complex interactions of molecular mediators of inflammation, oxidative stress, and immune responses. Transcriptomic and systems-level analyses have clearly shown the continuous activation of cytokine-mediated pathways in CF airway tissues, thereby highlighting the role of regulatory gene networks in the progression of CF (2; 3). Network-based bioinformatics tools facilitate the identification of central hub genes and pathways in disease interaction networks. Thus, this study uses network analysis to identify central molecular targets and pathways in CF, hoping to provide systems-level information that could help in developing additional therapeutic approaches to CF, independent of CFTR function.

Methodology

An integrated bioinformatics tool-based network analysis approach was developed to explore crucial genes, interaction modules, significant pathways, and possible therapeutic targets for cystic fibrosis. The current study combined disease gene mining, protein-protein interaction mapping, functional network construction, topological hub analysis, pathway enrichment analysis, and gene-drug sensitivity correlation. The current study used publicly available and peer-reviewed databases and bioinformatics tools to validate the results.

1. Seed gene identification – DisGeNET

The genes associated with cystic fibrosis (CF) were identified from the DisGeNET database, which is a carefully curated resource that provides gene-disease associations from the scientific literature and expert-reviewed sources. The search term used was "Cystic Fibrosis," and only genes that have been associated with humans were considered. Duplicates were removed before analysis (4).

Table 1: Seed gene identification – DisGeNET

Gene Symbol	Gene Description	Score	ei
CFTR	CF transmembrane conductance regulator	1.0	0.92981843575419
TGFB1	transforming growth factor beta 1	1.0	0.9473684210526315
SERPINA1	serpin family A member 1	0.95	0.9120879120879121
IL1B	interleukin 1 beta	0.9	0.868421052631579
TNF	tumor necrosis factor	0.9	0.8064516129032258
MPO	myeloperoxidase	0.9	0.8235294117647058
SCNN1A	sodium channel epithelial 1 subunit alpha	0.85	0.8777777777777778
SLC9A3	solute carrier family 9 member A3	0.85	1.0
MBL2	mannose-binding lectin 2	0.8	0.9629629629629629
HFE	homeostatic iron regulator	0.8	1.0
SCNN1B	sodium channel epithelial 1 subunit beta	0.8	0.9285714285714286
SLC6A14	solute carrier family 6 member 14	0.8	1.0
SLC26A9	solute carrier family 26 member 9	0.8	0.8518518518518519
ADRB2	adrenoceptor beta 2	0.8	1.0
AGER	advanced glycosylation end-product specific receptor	0.8	1.0
PTGS2	prostaglandin-endoperoxide synthase 2	0.8	0.8333333333333334

The genes were filtered based on Gene-Disease Association (GDA) scores (Score_gda ≥ 0.8), evidence index, number of publications supporting the gene (PMIDs), and evidence strength. High-confidence genes such as CFTR, TGFB1, SERPINA1, IL1B, TNF, MPO, MBL2, and HFE were identified as seed genes for further analysis.

This step ensured that only biologically validated and literature-supported genes were used as the seed for the construction of the network.

2. Protein-Protein Interaction (PPI) network construction – STRING

A protein-protein interaction (PPI) network was built using the STRING database (5) to examine functional interactions among the selected seed genes. STRING combines experimentally supported interactions, predictions, and pathway information.

High-confidence interaction scores were used to filter out false-positive interactions. A minimum interaction confidence score of 0.7 was applied. The network facilitated the visualization of interaction intensity, clustering, and the identification of connected regulatory modules. This systems-level analysis enabled the investigation of the interaction of inflammatory cytokines, immune receptors, and ion transport proteins in the CF molecular environment.

3. Functional network expansion – GeneMANIA

To extend the interaction network and identify other genes with functional relationships, GeneMANIA was used (6).

GeneMANIA combines co-expression information, physical interactions, pathway co-membership, genetic interactions, and shared protein domains. This step enabled the identification of secondary regulators and modifier genes that could potentially affect CFTR function, immune signaling, and oxidative stress responses.

Functional enrichment of co-expression and physical interaction data showed that these genes were functionally related in shared biological pathways, making the expanded network biologically relevant.

4. Topological analysis – NetworkAnalyst

The network topology measures of degree centrality were employed in NetworkAnalyst to identify the hub genes (7). The network centrality measures of degree centrality and interaction density were employed to identify the hub genes. Genes with high connectivity scores were regarded as important regulatory nodes owing to their role in network stability and information transfer. The hub genes of TNF, IL1B, and IL6 were selected based on their central location and interaction density in the inflammatory module.

5. Pathway enrichment analysis – Reactome

To identify the biological pathways that are significantly associated with the discovered hub genes, pathway enrichment analysis was performed with the Reactome database (8). The pathways with a false discovery rate (FDR) of < 0.05 were considered statistically significant. The enriched pathways were classified into the following: immune system signaling, cytokine-mediated signaling, oxidative stress response, ion transport regulation, and signal transduction. This analysis helped to assign meaning to the results of the network analysis and to connect the hub genes to biological processes underlying cystic fibrosis.

6. Gene expression and drug sensitivity correlation – GSCALite

For assessing possible therapeutic applications, gene expression & drug sensitivity correlations were investigated employing GSCALite (9). Correlation analysis was carried out between mRNA expression levels of hub genes and drug response datasets. Significant correlations (FDR ≤ 0.05) were determined,

where:

- Positive correlation suggested possible drug resistance related to high gene expression.
- Negative correlation suggested higher drug sensitivity related to high gene expression.

This combined analysis facilitated the detection of possible predictive biomarkers and therapeutic targets in the CF molecular network.

Results

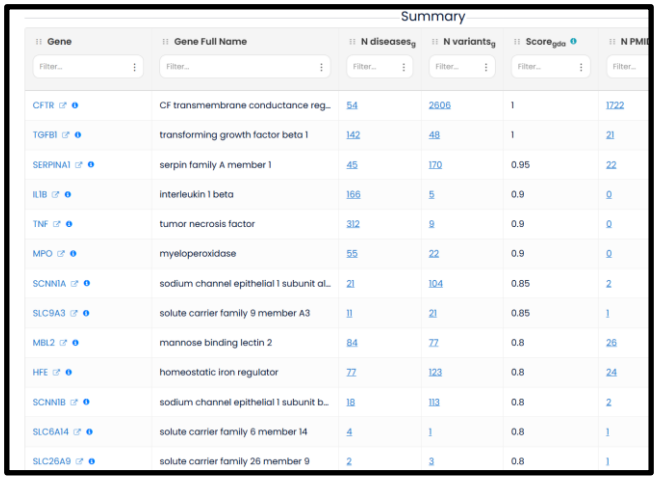
1. DisGeNET – Seed gene selection

A total of 16 high-confidence genes were selected. Seed genes were selected using the DisGeNET database based on Gene-Disease Association (GDA) scores and evidence. Genes were ranked based on high association scores (Score_gda ≥ 0.8), the number of associated diseases, the number of variants, and literature evidence (PMIDs). Several high-confidence genes were selected using the analysis, including CFTR, TGFB1, SERPINA1, IL1B, TNF, MPO, MBL2, and HFE.

Among these:

- CFTR and TGFB1 have the highest association score (Score_gda = 1), reflecting strong curated evidence.
- IL1B and TNF have high disease association counts, reflecting their role in inflammatory and immune mechanisms.
- MPO and HFE are involved in hematological and metabolic regulatory mechanisms.
- Cytokine and signaling genes (TGFB1, IL1B, TNF) reflect the enrichment of immune and inflammatory pathways.

The high GDA scores and literature evidence reflect the strength of the seed genes selected. These genes were thus selected for downstream protein-protein interaction and network analyses.



Summary

Gene	Gene Full Name	N diseases$_g$	N variants$_g$	Score$_{gda}$	N PMI
CFTR	CF transmembrane conductance reg...	54	2606	1	1722
TGFB1	transforming growth factor beta 1	142	48	1	21
SERPINA1	serpin family A member 1	45	170	0.95	22
IL1B	interleukin 1 beta	166	5	0.9	0
TNF	tumor necrosis factor	312	9	0.9	0
MPO	myeloperoxidase	55	22	0.9	0
SCNN1A	sodium channel epithelial 1 subunit al...	21	104	0.85	2
SLC9A3	solute carrier family 9 member A3	11	21	0.85	1
MBL2	mannose binding lectin 2	84	77	0.8	26
HFE	homeostatic iron regulator	77	123	0.8	24
SCNN1B	sodium channel epithelial 1 subunit b...	18	113	0.8	2
SLC6A14	solute carrier family 6 member 14	4	1	0.8	1
SLC26A9	solute carrier family 26 member 9	2	3	0.8	1

Figure 1: Seed genes with higher confidence extracted from the DisGeNET web tool (https://disgenet.com/)

2. STRING – Protein-Protein Interaction (PPI) network analysis

Protein-protein interaction analysis was carried out using the STRING database to assess functional association between the identified seed genes. The PPI network analysis showed a strongly interconnected interaction map, with several hub nodes such as IL1B, TNF, TGFB1, IL6, TLR4, MPO, and CFTR. These genes showed strong interaction patterns, indicating their involvement in common regulatory and signaling pathways.

The PPI network showed:

- A strongly clustered group of inflammatory cytokines (IL1B, TNF, IL6).
- Strong interactions between immune receptors (TLR2, TLR4, TLR5) and signaling molecules.
- Strong associations between oxidative stress genes (MPO, HMOX1) and immune regulators.
- Strong peripheral associations between transport and channel genes (SLC family members, SCNN1A/B/G).

The strongly interconnected map suggests a coordinated role in immune regulation, inflammation, and cellular stress responses. The STRING analysis clearly shows that the identified genes constitute a biologically valid network for further analysis of hub genes and pathways.

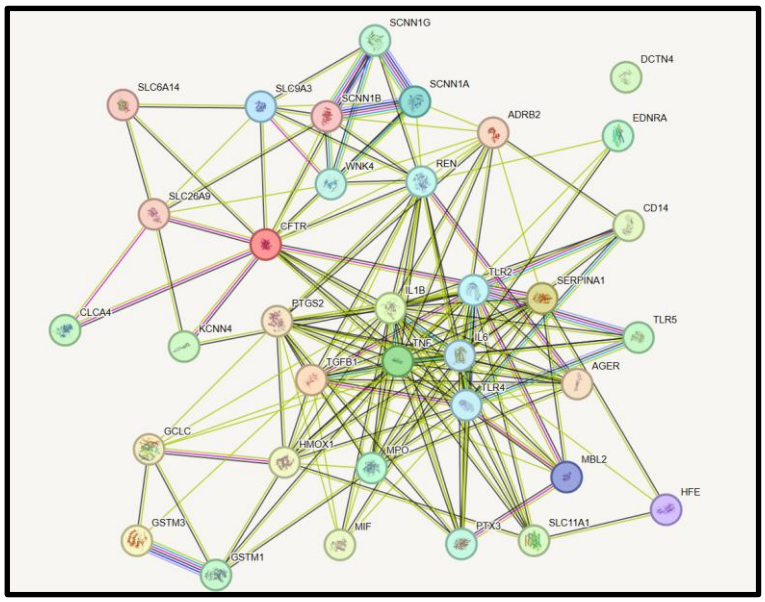


Figure 2: Protein–Protein Interaction Network of CF-associated genes constructed using STRING

3. GeneMANIA – Functional network expansion

Functional network enlargement was carried out using GeneMANIA to find other genes that are functionally associated with the chosen seed genes based on co-expression, physical interactions, pathway associations, and genetic interactions.

The functional network analysis showed a high degree of functional connectivity among the important inflammatory and immune response genes, such as IL1B, TNF, TGFB1, IL6, CFTR, MPO, and SERPINA1, and newly identified genes such as PTGS2, TLR3, TLR5, GSTM1, SLC family members, and SCNN1 gene variants.

The analysis showed:

- Immune and inflammatory signaling interactions are enriched.
- Oxidative stress-related genes (HMOX1, GSTM1, GSTM3) are integrated.
- Functional clustering of transporters and ion channel genes (SLC6A14, SLC9A3, SCNN1A/B/G) is observed.
- Cytokine signaling and innate immune receptors (TLRs) interact.

The preponderance of co-expression and physical interaction network links indicates a coordinated regulation of genes in common biological pathways. In summary, the GeneMANIA analysis further refines the functional

network by identifying other functionally relevant genes, which are useful in downstream hub gene identification and pathway enrichment analysis.

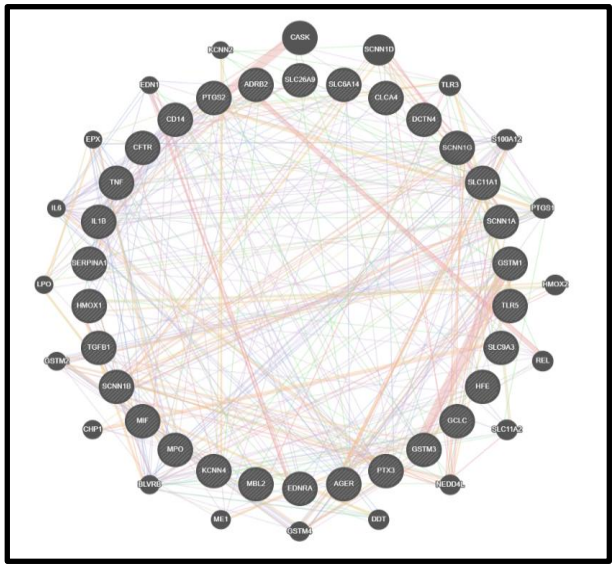


Figure 3: PPI Network of existing (seed genes) as well as additional genes involved in various pathways (GeneMANIA n.d.).

4. NetworkAnalyst- Hub gene selection

Topological analysis of the enlarged interaction network was carried out using NetworkAnalyst to identify the most important hub genes, depending on their connectivity and centrality scores.

Results from the analysis showed that TNF, IL1B, IL6, and SERPINA1 were the central hub genes with high degree scores, indicating a high level of interaction density in the network. Among these, TNF was shown to be the most central hub gene with a high level of connectivity, indicating a dense interaction cluster with multiple downstream mediators. IL1B and IL6 also showed high centrality scores, indicating a coordinated role in the inflammatory signaling pathways.

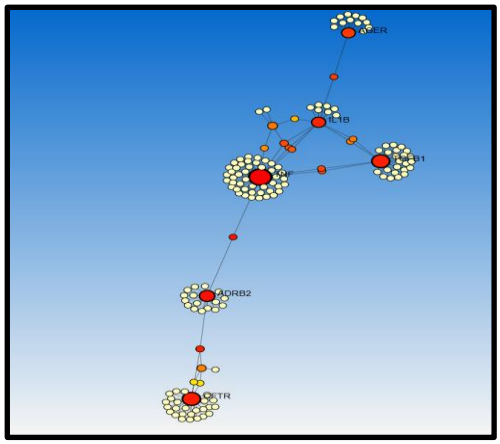


Figure 4: Network of genes with central connections and interactions with central nodes (Hub Genes), in a topological pattern (NetworkAnalyst n.d.)

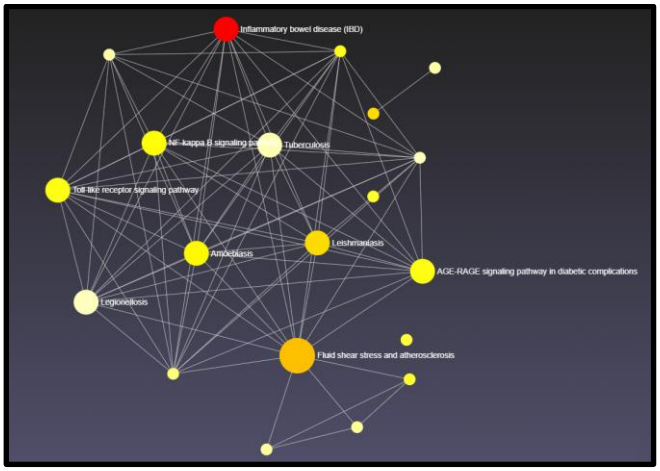


Figure 5: Network of genes with central connections and interactions with central nodes (Hub Genes), in a topological pattern (NetworkAnalyst n.d.).

Peripheral but connected hubs such as CFTR, ADRB2, and AGER were shown to form smaller functional modules that are connected to the central inflammatory hub. The existence of sub-clusters in the network indicates a modular structure, with the inflammatory signaling pathway being the central regulatory module.

The identification of TNF, IL1B, and IL6 as central hub genes indicates the dominance of inflammatory and cytokine-mediated pathways in the network structure. These hub genes can be considered as potential key targets for further pathway and therapeutic analysis.

5. Reactome – Pathway enrichment analysis

Pathway enrichment analysis was performed using the Reactome Consortium knowledgebase to identify biological pathways significantly associated with the identified hub genes (TNF, IL1B, IL6, SERPINA1, CFTR, TGFB1). Pathways with False Discovery Rate (FDR) < 0.05 were considered statistically significant.

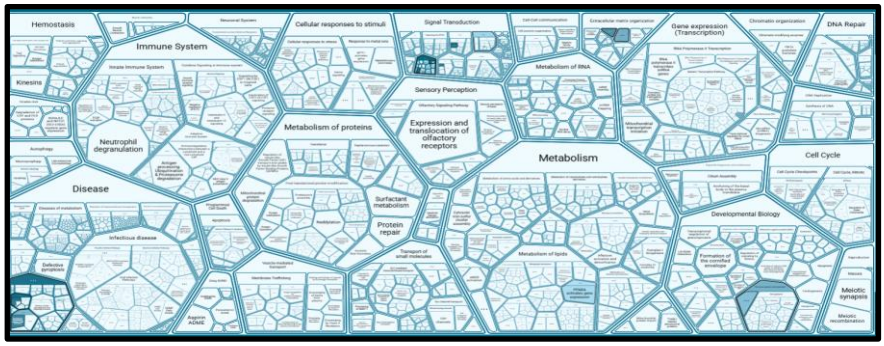


Figure 6: Map of Pathways involved in Upregulation of Hub Genes that lead to CF- (Highlighted are the Major Pathways) (Reactome n.d.)

Biological Interpretation of CF:

- TNF, IL1B, and IL6 were highly enriched in cytokine-mediated inflammatory signaling pathways.
- Neutrophil degranulation pathways represent chronic neutrophilic inflammation in CF lungs.
- Toll-like receptor signaling (TLR2/TLR4) suggests innate immune activation in response to chronic bacterial colonization.

- Oxidative stress pathways (MPO, HMOX1) confirm the role of reactive oxygen species-mediated epithelial injury.
- Ion transport pathways (CFTR, SCNN1A/B) verify dysfunctional chloride and sodium transport as the molecular basis of CF disease.
- TGFB1 enrichment suggests a role in airway remodeling and fibrosis.

These findings align with CF being a chronic inflammatory and epithelial transport disorder.

Table 2: Pathway annotations

Major Pathway	Description	Relevance to CF	Sources
Immune System	Broad immune activation pathways	Chronic airway inflammation	1, 2
Cytokine Signaling	IL1B, IL6, TNF-mediated cascades	Persistent inflammatory response	2, 3
TNF Signaling	NF-κB–mediated transcription	Mucus hypersecretion & inflammation	10
Interleukin-1 Signaling	Pro-inflammatory activation	Neutrophil recruitment	2, 3
Neutrophil Degranulation	Release of proteases & ROS	Tissue damage in CF lungs	1, 11
Toll-like Receptor Cascades	Bacterial recognition pathways	Response to Pseudomonas infection	2, 3
Oxidative Stress Response	ROS generation & detox pathways	Airway epithelial injury	3, 12
Ion Transport	CFTR and ENaC regulation	Core defect in CF	1, 10
Extracellular Matrix Organization	Fibrosis and tissue remodeling	Progressive lung damage	2, 3

6. GSCALite – Gene expressions and drug sensitivity analysis

Correlations between gene expression and drug sensitivity were investigated using GSCALite to evaluate the possible therapeutic relevance of hub genes in inflammatory and immune regulation.

Significant correlations (FDR ≤ 0.05) were:

- Positive correlation → Increased gene expression with decreased drug sensitivity (potential resistance)
- Negative correlation → Increased gene expression with increased drug sensitivity
- In the CF scenario:
- Increased TNF and IL1B gene expression could predict treatment response to anti-inflammatory or cytokine-based therapies.
- Increased CFTR gene expression could correlate with treatment response to CFTR modulators.
- Increased oxidative stress gene expression (MPO, HMOX1) could affect treatment response to antioxidant therapies.

This indicates that hub genes might be used as predictive biomarkers for personalized CF treatment. Drug sensitivity analysis was exploratory and based on available transcriptomic-drug response datasets.

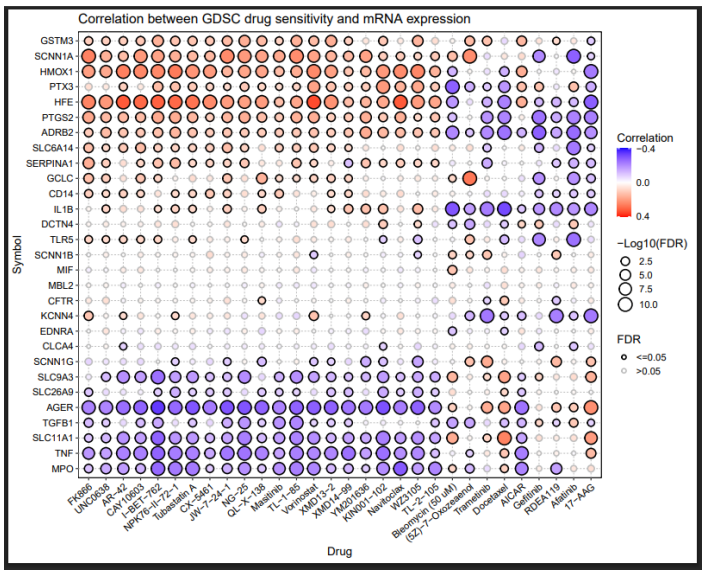


Figure 7: Bubble plot shows significant gene–drug correlations in CF

Discussion

This network analysis shows that cystic fibrosis (CF) pathogenesis is not solely due to primary CFTR dysfunction but also involves complex inflammatory signaling pathways. While CFTR mutations are the initial cause of epithelial ion transport dysfunction, the network topology analysis showed that TNF, IL1B, and IL6 are key hub genes, emphasizing the central role of cytokine-mediated pathways in maintaining chronic airway inflammation. The persistent activation of these inflammatory mediators has been shown in CF transcriptomic analyses and is associated with neutrophilic lung injury and disease severity (2).

Pathway enrichment analysis also showed the significant role of Interleukin-1 signaling, neutrophil degranulation, and oxidative stress pathways, which provide a mechanistic insight into how exaggerated immune responses and redox dysregulation contribute to progressive epithelial injury (12). These observations are consistent with systems-level analyses suggesting that CFTR is part of a larger regulatory network that modulates immune responses and cellular homeostasis (3).

Taken together, the set of highly interconnected hub genes suggests additional therapeutic targets that could serve as adjuncts to CFTR modulator therapies. Nevertheless, experimental confirmation is necessary to validate the biological relevance of the proposed interaction network.

Conclusion

In summary, this bioinformatics-informed network analysis offers a systems-level view of cystic fibrosis pathogenesis. In addition to CFTR dysfunction, this study emphasizes the essential role of inflammatory and immune regulatory networks, with TNF, IL1B, and IL6 being key hub genes. The enrichment of Interleukin-1 signaling, neutrophil degranulation, and oxidative stress pathways emphasizes the role of chronic inflammatory activation in disease progression. These results support the notion that cystic fibrosis is a complex inflammatory channelopathy, rather than a simple ion transport disease. The hub genes and pathways identified in this study could potentially be used as targets for therapy, in addition to CFTR modulator therapies. Experimental validation and multi-omics analysis are needed to support these bioinformatics predictions and facilitate the development of precision medicine strategies for cystic fibrosis.

References

1. Rafeeq, M. M., & Murad, H. A. (2017). Cystic fibrosis: Current therapeutic targets and future approaches. *Journal of Translational Medicine, 15*, 84. https://doi.org/10.1186/s12967-017-1193-9
2. Kormann, M. S. D., Dewerth, A., Eichner, F., Fischer, B., et al. (2017). Transcriptomic profile of cystic fibrosis patients identifies gene networks influencing disease severity. *PLOS ONE, 12*(8), e0183526. https://doi.org/10.1371/journal.pone.0183526
3. Najm, M., Martignetti, L., Cornet, M., Kelly-Aubert, M., Sermet, I., Calzone, L., & Stoven, V. (2024). From CFTR to a cystic fibrosis signalling network: A systems biology approach to study cystic fibrosis. *BMC Genomics, 25*, 10752. https://doi.org/10.1186/s12864-024-10752-x
4. DisGeNET Consortium. (2020). Gene–disease association platform. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw943
5. Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S. (2021). The STRING database in 2021: Customizable protein–protein networks and functional characterization of user-uploaded gene sets. *Nucleic Acids Research, 49*(D1), D605–D612. https://doi.org/10.1093/nar/gkaa1074
6. Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P. (2010). The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research, 38*(Suppl 2), W214–W220. https://doi.org/10.1093/nar/gkq537
7. NetworkAnalyst Team. (2020). NetworkAnalyst 3.0: Visual analytics platform for gene expression and network analysis. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkz240
8. Reactome Consortium. (2022). Reactome pathway knowledgebase 2022. *Nucleic Acids Research, 50*(D1), D687–D692. https://doi.org/10.1093/nar/gkab1028
9. Liu, C.-J., Hu, F.-F., Xia, M.-X., Han, L., Zhang, Q., & Guo, A.-Y. (2018). GSCALite: A web server for gene set cancer analysis. *Bioinformatics, 34*(21), 3771–3772. https://doi.org/10.1093/bioinformatics/bty411
10. Lim, S. H., Legere, E.-A., Snider, J., & Stagljar, I. (2018). Recent progress in CFTR interactome mapping and its importance for cystic fibrosis. *Frontiers in Pharmacology, 8*, 997. https://doi.org/10.3389/fphar.2017.00997
11. O'Connor, S., *et al*. (2022). Network analysis to identify multi-omic correlations in cystic fibrosis airways of children. *Frontiers in Cellular and Infection Microbiology, 12*, 805170. https://doi.org/10.3389/fcimb.2022.805170
12. Strub, M. D., *et al.* (2020). Transcriptomic and proteostasis networks of CFTR and implications for cystic fibrosis. *Genes, 11*(5), 546. https://doi.org/10.3390/genes11050546