



OPEN

Transcriptomic analysis delineates potential signature genes and miRNAs associated with the pathogenesis of asthma

Prithvi Singh^{1,5}, Archana Sharma^{2,5}, Rishabh Jha¹, Shweta Arora², Rafiq Ahmad³, Arshad Husain Rahmani⁴, Saleh A. Almatroodi⁴, Ravins Dohare^{1✉} & Mansoor Ali Syed^{2✉}

Asthma is a multifarious disease affecting several million people around the world. It has a heterogeneous risk architecture inclusive of both genetic and environmental factors. This heterogeneity can be utilised to identify differentially expressed biomarkers of the disease, which may ultimately aid in the development of more localized and molecularly targeted therapies. In this respect, our study complies with meta-analysis of microarray datasets containing mRNA expression profiles of both asthmatic and control patients, to identify the critical Differentially Expressed Genes (DEGs) involved in the pathogenesis of asthma. We found a total of 30 DEGs out of which 13 were involved in the pathway and functional enrichment analysis. Moreover, 5 DEGs were identified as the hub genes by network centrality-based analysis. Most hub genes were involved in protease/antiprotease pathways. Also, 26 miRNAs and 20 TFs having an association with these hub genes were found to be intricately in a 3-node miRNA Feed-Forward Loop. Out of these, miR-34b and miR-449c were identified as the key miRNAs regulating the expression of SERPINB2 gene and SMAD4 transcription factor. Thus, our study is suggestive of certain miRNAs and unexplored pathways which may pave a way to unravel critical therapeutic targets in asthma.

Despite recent advances in anti-asthmatic therapeutics, asthma is still a major global health concern. According to the latest Global Asthma Report 2018, over 339 million people are affected globally by asthma leading to more than 1000 deaths per day (<https://globalasthmareport.org>). At least 2 billion individuals worldwide have exposure to the contaminated smoke of biomass fuel, usually burned inadequately in fire replaces or poorly ventilated indoor stoves. A billion individuals inhale polluted outdoor air, and a billion have exposure to tobacco smoke¹. Asthma is a chronic disorder of airways, which is characterized by inflammation resulting from a complex interplay of diverse pathways in numerous type of cells including epithelial cells, smooth muscle cells, neutrophils, mast cells, eosinophils, T cells, and B cells^{2,3}. Cellular and molecular processes in asthma are relatively unknown due to heterogeneity in genetic, clinical and treatment responses in asthma subjects. Airway re-modelling takes places in severe asthma characterized by goblet cell metaplasia, epithelial damage, sub-epithelial fibrosis, basement membrane thickening, and escalated airway smooth muscle^{4,5}. Based on phenotypes, there can be non-allergic asthma, allergic asthma, asthma with obesity, adult-onset asthma, and asthma with persistent airflow limitation^{6,7} whereas based on the severity it could be mild, moderate or severe. Molecular phenotyping of diseased tissues helps to identify biomarkers of asthma which are essential for the development of more localized and molecularly targeted therapies based on the heterogeneity of patients⁸.

There are two subgroups of asthma based on the expression of Th2 associated genes. One subgroup expresses a higher level of Th2 associated genes. While another subgroup which lacks few Th2 associated genes in airways epithelial cells is termed as non-Th2 associated. This non-Th2 asthma group may be low-Th2 asthma/ Th1 driven

¹Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi 110025, India. ²Translational Research Lab, Department of Biotechnology, Faculty of Natural Sciences, Jamia Millia Islamia, New Delhi 110025, India. ³Centre for Nanoscience and Nanotechnology, Jamia Millia Islamia, New Delhi 110025, India. ⁴Department of Medical Laboratories, College of Applied Medical Sciences, Qassim University, Buraydah 51452, Saudi Arabia. ⁵These authors contributed equally: Prithvi Singh and Archana Sharma. ✉email: ravinsdohare@gmail.com; smansoor@jmi.ac.in

asthma. In Th2 group asthma patients, CD4⁺ T cells produce IL-4, IL-5, and IL-13 (Th2 cells), which have been reported in bronchoalveolar lavage and airway autopsies. There has been an anticipation that asthma is resulted due to an inflammatory response stimulated by Th2 cells. Increment of Th2 lymphocytes in the airways of allergic asthmatic patients occurs after antigen challenge. These are the cytokines which have been usually reported in mild or asymptomatic asthma⁹. IL-13 and IL-4 are the potent activators of antibody production by B cells especially IgE, while IL-5 is crucial for eosinophil maturation and differentiation. Although there are no unanimously accepted biomarkers for Th2 asthma group, they are characterized by the atopy, eosinophilic inflammation, and a good clinical response to inhaled corticosteroids (ICSs)¹⁰. The non-Th2 asthma group is usually less responsive to ICSs. The non-Th2 asthma group includes asthma driven by smoking, pollutants, viral/ bacterial infections or obesity, late-onset asthma, etc. Low Th2 asthma can be due to many cellular signalling pathways, including neutrophilic inflammation by dysregulation of the Th17 pathway, oxidative stress, Th1 mediated processes which may include a Th2 component also. Th1 cells were identified in many non-Th2 patients' airways which play a role in the pathogenesis of the disorder. Study of IFN- γ production in the patients indicated that Th1 cells in asthmatics are not allergen-specific and these may be either bystander in airways or must be playing some regulatory capacity role^{11,12}.

The primary objective of the research is to pinpoint the Differentially Expressed Genes (DEGs) between asthmatics and controls from mRNA expression profiles of epithelial cells lining Nasal and Bronchial airway from two different studies to strengthen the statistical analysis. Pathway and Gene Ontology (GO) term enrichment analysis of these DEGs helps to uncover which pathways are dysregulated in asthma. Protein-Protein Interaction (PPI) network was also formed which helped in deciphering the underlying molecular mechanism in the pathophysiology of the disease. Subsequently, hub genes were elucidated from the PPI network by taking into consideration some of the vital network centrality measures. Finally, driver microRNAs (miRNAs) and Transcription Factors (TFs) associated with these asthma-associated DEGs are identified via Feed-Forward Loop (FFL).

In accordance with graph theory, the basic information of the PPI network is provided by its centralities. The blend of PPI centrality measures and important biological knowledge acts as a mean for identifying the biological mechanism of species¹³. Node degree centrality, closeness centrality, betweenness centrality, stress centrality, and EPC are the essential measures in network theory that have been used for detecting effective connections in biological networks. Recently, centrality-based measures for hub genes identification have been applied to many networks such as PPI networks, gene co-expression networks, and cancer metabolism networks¹⁴. Ernesto Estrada et al.¹⁵ showed that subgraph-based six centrality measures overpowered classic measures for identifying vital proteins in a PPI network. Jeong et al.¹⁶ exhibited that nodes with high degree centrality in a yeast PPI network correspond to crucial proteins.

The DEG biomarkers identified using computational approach are comprehensively regulated by TFs and miRNAs. TFs are cis-regulating molecules which act on gene's promoter region thereby regulating gene transcription. On the other hand, miRNAs are the small (20–22 nucleotide) non-coding RNA molecules which regulate gene expression post-transcriptionally. miRNAs exert their control by acting post-transcriptionally by binding in 3' UTR of the gene. They are known to exert epigenetic control over various cellular process like cell division, differentiation, apoptosis, disease progression, etc.¹⁷. TFs and miRNAs act in an intricate regulatory network which is tightly coupled to form a common regulatory logic for gene regulation¹⁸. In our study, we utilized microarray-based mRNA expression profiles of bronchial and nasal airway samples of asthmatic and control patients. Both nasal and bronchial airway epithelium is composed of ciliated, basal, and secretory epithelial cells¹⁹. The significant regulatory networks identified from microarray-based expression profiling of bronchial and nasal airway can be combined to form a transcriptional/ post-transcriptional FFL. In this study, we pursued a regulatory network-based approach to identify critical miRNAs, DEGs and TFs which form asthma specific 3-node miRNA FFL network. Study of these regulatory FFLs may lead to the identification of specific biological events and mechanisms which are responsible for determining cell fate in asthma and other allergic diseases. Identification of critical regulatory genes, TFs, and miRNAs targeting the Th2 inflammatory pathway or the altered gene expression of the bronchial or nasal airway in asthma patients would be a valuable tool in elucidating pathogenesis and therapeutic strategies in asthma.

Results

Microarray data extraction and preprocessing. Based on the inclusion and exclusion criteria specified in the methods section, microarray gene expression profiles with accession numbers GSE41861 and GSE41862 were downloaded. GSE41861 contains expression data from Bronchial Samples (BS) and Nasal Samples (NS) of 91 asthmatic patients and 47 controls, respectively. And, GSE41862 contains expression data from BS of 95 asthmatic patients and 21 controls, respectively. Information related to these datasets like GEO accession number, platform type, number of samples, type of study, and species is shown in Supplementary Table S1, respectively. The intensity scatter plots for visualizing the variance in datasets GSE41861 and GSE41862 are shown in Supplementary Figure S1.

Meta-analysis and identification of DEGs. A total of 30 DEGs were yielded after meta-analysis of microarray datasets involving asthmatics vs controls and applying the statistical threshold of *BH-p*-value and \log_2 (fold change) respectively. Figure 1 shows the Principal Component Analysis (PCA) plot (left panel) representing the variation and clusters of 30 asthma-associated DEGs across 2 dimensions, while the scree plot (right panel) represents the percentage of variances accounted for by the principal components, respectively. Among the total DEGs, 18 were upregulated and 12 were downregulated based on the \log_2 (fold change) criterion. Annotated heatmap of top 10 up and downregulated DEGs is shown in Fig. 2.

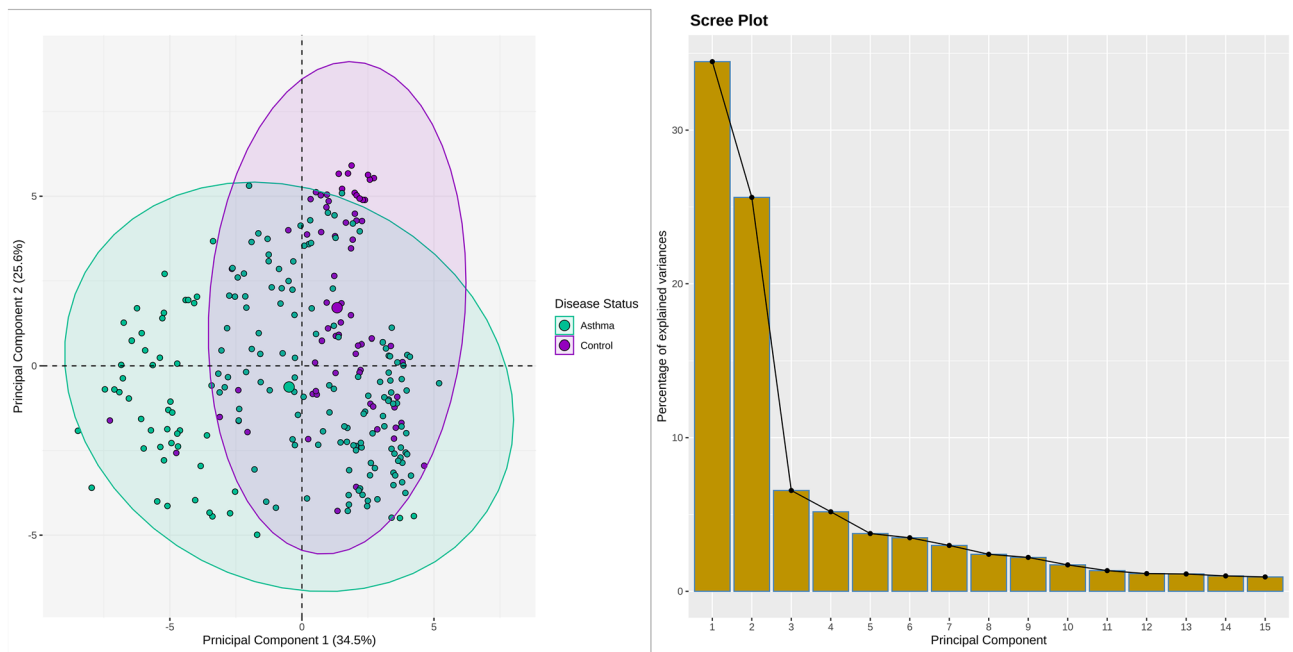


Figure 1. PCA plot in the left panel represents the variation in the expression data between controls and asthmatics. Each point in the plot shows the overall expression value of 30 screened DEGs. The color of each point represents the disease status: magenta for controls and green for asthmatics. The percentage of total variation which is accounted for by the 1st and 2nd principal components are shown on the x and y axes respectively. Scree plot in the right panel shows the percentage of explained variances captured by their corresponding principal components.

Pathway & GO term enrichment analysis. A total of 13 DEGs out of 30 were involved in significantly enriched pathways and GO terms. Nine DEGs (MUC5B, POSTN, SERPINB2, CPA3, GRP, LTF, MS4A2, TPSAB1, and WIF1) were involved in a total of 17 significantly enriched pathways (BH - p -value < 0.05) with peptide hormone metabolism being the most significant pathway (BH - p -value = 0.006547577) and ROS, RNS production in response to bacteria being the least significant pathway (BH - p -value = 0.04979756). A circo plot showing the association of 9 DEGs with 17 significantly enriched pathways is shown in Fig. 3. Six DEGs (SERPINB10, SERPINB4, SERPINB2, LTF, CST1, and CST4) were involved in significantly enriched GO-Biological Process (BP) term: negative regulation of endopeptidase activity (GO:0010951) and GO-Molecular Function (MF) terms: endopeptidase inhibitor activity (GO:0004866), serine-type endopeptidase inhibitor activity (GO:0004867), and cysteine-type endopeptidase inhibitor activity (GO:0004869) respectively (Supplementary Table S2).

PPI network analysis and hub gene identification. Unique 13 DEGs (9 upregulated and 4 downregulated DEGs) obtained after enrichment analysis were further used for the PPI network construction and analysis using Cytoscape. Both up and downregulated PPI networks are shown in Supplementary Figures S2 and S3. The topological properties of up and downregulated PPI networks are shown in Supplementary Table S3 respectively. Top n functionally enriched DEGs ($n=9$ for upregulated & $n=4$ for downregulated) were ranked based on the centrality measures. A total of 5 intersecting hub genes (3 upregulated, i.e., CST4, SERPINB2, SERPINB4 and 2 downregulated, i.e., LTF, MUC5B) were obtained and shown by the Venn plots in Fig. 4A and B respectively. Violin plots showing the expression distribution of these hub genes in controls and asthmatics are shown in Fig. 5 respectively. Among all the hub genes, CST4 and LTF were most upregulated and downregulated across asthmatic samples.

Asthma-specific miRNA FFL regulatory network. The 3-node miRNA FFL involved a total of 51 unique nodes and 197 edges (Fig. 6). Among these edges, 32 belonged to miRNA-gene pairs, 21 to TF-gene pairs, and 144 to miRNA-TF pairs as shown in Table 1. Amongst the 51 nodes, 5 belonged to asthma-associated hub genes, 26 to asthma-associated miRNAs and 20 to asthma-associated human TFs. Different topological components, including the distribution of node degree, average clustering coefficient, topological coefficient, betweenness centrality, closeness centrality, and shortest path length for the miRNA FFL network are shown in Fig. 7A–F respectively. The node degree distribution for miRNAs, TFs, and genes in the miRNA-FFL was right-skewed, depicting that majority of nodes had a low degree, with only a minor section of nodes having a high degree. Most strongly connected network motif or subnetwork in our FFL contained one TF (SMAD4), one hub gene (SERPINB2), and two miRNAs (miR-449c & miR-34b-5p).

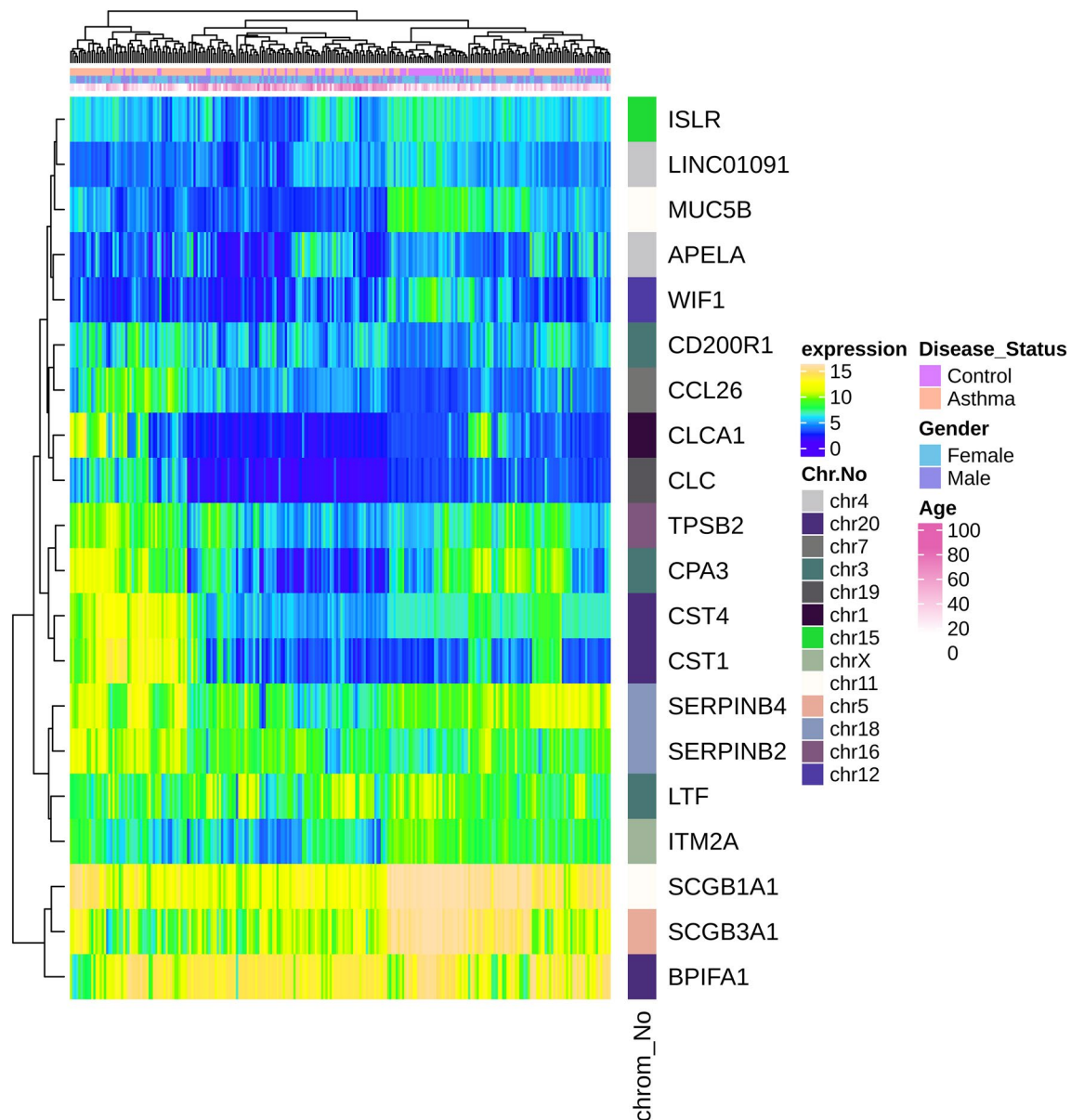


Figure 2. Heatmap plot of top 10 up and downregulated DEGs in asthma. Expression values for each DEG (row) are normalized across all the samples (columns). Hierarchical clustering using Euclidean distances was employed for both rows and columns with cluster dendrograms presented along the left and top sides of the plot. The categorical annotation bars (above the heatmap) show the column annotations for age, gender and disease status. The location of each gene on its respective chromosome was shown in the right bar (color bands) as row annotation.

miRNA-gene and miRNA-TF repression. A total of 26 miRNAs targeted the five hub genes. SERPINB2 was repressed by maximum miRNAs (i.e.,13). Also, both miR-34b-5p and miR-449c-5p targeted highest number of genes (i.e.,3). miRNA-TF interaction had 144 edges between 26 miRNAs and 20 TFs respectively. TCF4 was repressed by the highest number of miRNAs (i.e., 17). miR-423-5p repressed highest number of TFs (i.e.,12).

TF-gene regulation. A total of 20 TFs regulated the five hub genes where both LTF and MUC5B were regulated by a maximum number of TFs (i.e., 7). Also, NR2F1 targeted the highest number of hub genes (i.e., 2).

Discussion

Transcriptional regulatory networks are the key functional motif networks which interplay in various diseases including asthma leading to a molecular pathological outcome. During recent years, databases of interaction prediction, verification of experimental methods and confirmation of loops have been continually emerging²⁰. The integrated results of five important PPI-based centrality measures (EPC, degree, stress, betweenness and closeness) leading to vital hub genes resolve the outcomes which were inconsistent as provided by different

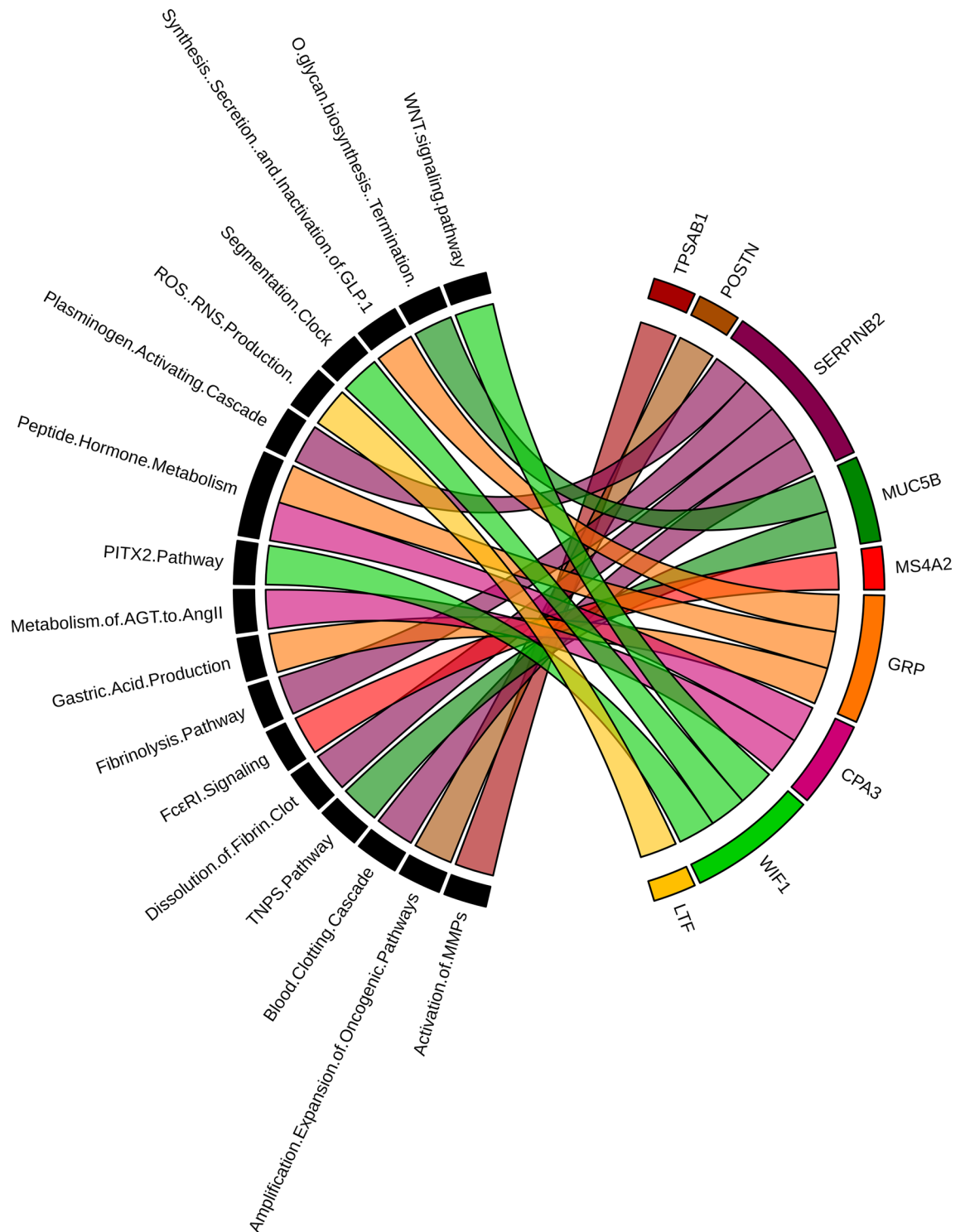


Figure 3. Circos plot representation of significantly enriched pathways linked with 9 asthma-associated DEGs. Outside the circle, 17 significantly enriched pathways (on the left) and 9 dysregulated genes (on the right) are indicated. Each DEG is denoted by a unique color band and the undirected colored edge inside the circle represents the association of a particular gene with their respective pathway(s).

other methods²¹. In our study, we identified 30 asthma-linked DEGs after the meta-analysis of microarray datasets and deduced 13 functionally enriched DEGs which formed the two backbone PPI networks. From these two networks, 5 hub genes were found based on their degrees and other centrality measures. The upregulated genes being CST4, SERPINB2 and SERPINB4, while downregulated genes being LTF and MUC5B, respectively. Asthma-associated miRNAs and TFs having a relationship with these hub genes were also discovered. In our study, we utilized key transcriptional level interaction i.e. miRNA-gene repression, miRNA-TF repression, and

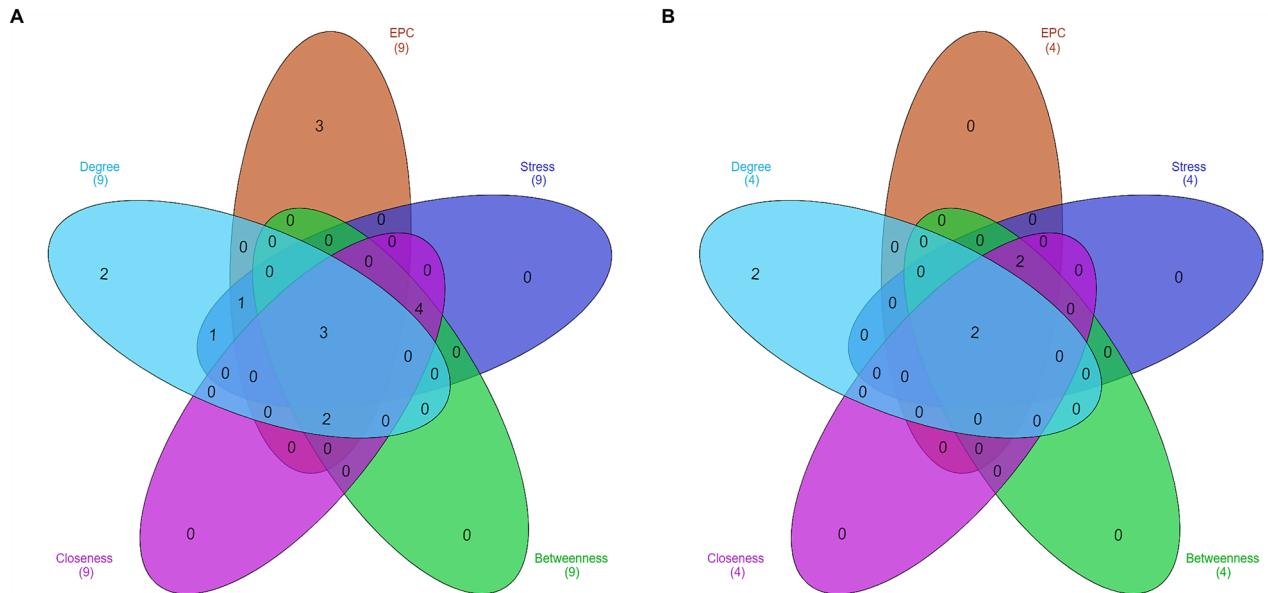


Figure 4. Venn plots showing the significant hub genes obtained from (A) upregulated and (B) downregulated PPI networks, respectively. Areas with different colors correspond to different centrality measures. The cross-intersection area shows the hub genes in both networks. A total of 5 hub genes were obtained for the up and downregulated PPI networks.

TF-gene regulation for the construction of miRNA centred feed-forward loop which plays a critical role in asthma (Fig. 6). The results of our study identified two key miRNAs, i.e., miR-34b and miR-449c which regulated the SERPINB2 gene and SMAD4 transcription factor, being the part of FFL.

Most of the hub genes observed in our study were involved in protease/antiprotease pathways (SERPINB2, SERPINB4, CST4). These upregulated hub genes are also indicative that neutrophilic inflammation in asthma may be driven by the activated T cells. Cystatin S (CST4), which is an epithelial gene, was found to be a differentially expressed hub gene in our study as also revealed by other research groups in different study cohort and methods^{22,23}. Cystatins have recognized roles in inflammation where they act as potent competitive inhibitors of cysteine proteinases (cathepsins) being present in all biological fluids²⁴. Induction of synthesis of Tumor necrosis factor- α (TNF- α) and IL-10 are among the potent immunomodulatory roles of cystatins. Another hub gene Mucin-5B (MUC5B) is a marker of secretory cells which was downregulated in asthmatic subjects in our study. MUC5B is expressed predominately in the mucous glands. Although our study datasets didn't define the severity level of asthma, epithelial MUC5B repression was observed in mild-moderate asthma by other studies²³. Changes in the expression of this gene indirectly reflect the airway remodelling in case of chronic asthma patients²⁵.

Lactoferrin or lactotransferrin (LTF) serves as an anti-inflammatory factor and an immunomodulator. LTF which is an iron-binding glycoprotein has been shown upregulated during asthma development in our study as also reported by other studies^{26,27}. A report by Lourdes Fernández-Delgado et al. suggested a new mechanism of LTF release by human neutrophils on activation by galectin-3 and claimed the pro-inflammatory role of LTF in allergic asthma²⁸. LTF could reduce airway inflammation in the pollen-induced allergic mouse model and its allergen mediated release could be correlated to the severity of asthma symptoms^{28,29}. Thus, being a hub gene also as elucidated in our study, LTF can be suggested as a putative tool to track the progression or the drug efficacy in allergic asthma patients.

Serine protease inhibitors (SERPIN) genes like SERPINB2 and SERPINB4 have an important role in asthma pathogenesis. SERPINB2 or Serpin B2 prevents plasmin activation by inhibiting plasminogen activators. Plasmin can degrade ECM directly by removal of glycoproteins or indirectly by activation of metalloproteinases³⁰. Therefore, by inhibition of SERPINB2, plasmin availability for extracellular matrix turnover increases and reduces airway remodeling. Increased SERPINB2 in airway epithelial cells has been described as a biomarker gene signature for Th2-mediated inflammation in asthma³¹. It is induced by Th2 cytokine interleukin (IL-13) which also acts as the central regulator of goblet cell metaplasia³². An increase in mucous cell number (mucous cell metaplasia) and a decrease in ciliated cells number has been linked with epithelial cells stimulation by IL-13³³. Obstruction of airways by mucus is a key contributor to fatal asthma. As it has been known that changes in airway epithelial cell differentiation, driven in part by IL-13 are essential in asthma, miRNAs are known to regulate cell differentiation in various systems and could contribute to epithelial abnormalities in asthma. SERPINB4 or SCCA2 gene is a potent inhibitor of apoptosis. It can inhibit apoptotic serine proteinases, such as cathepsin G and mast cell chymase which results in the persistence of memory Th2 cells that can give rise to inflammatory lung disease like asthma³⁴. It was identified as an upregulated hub gene in our study as also described by other

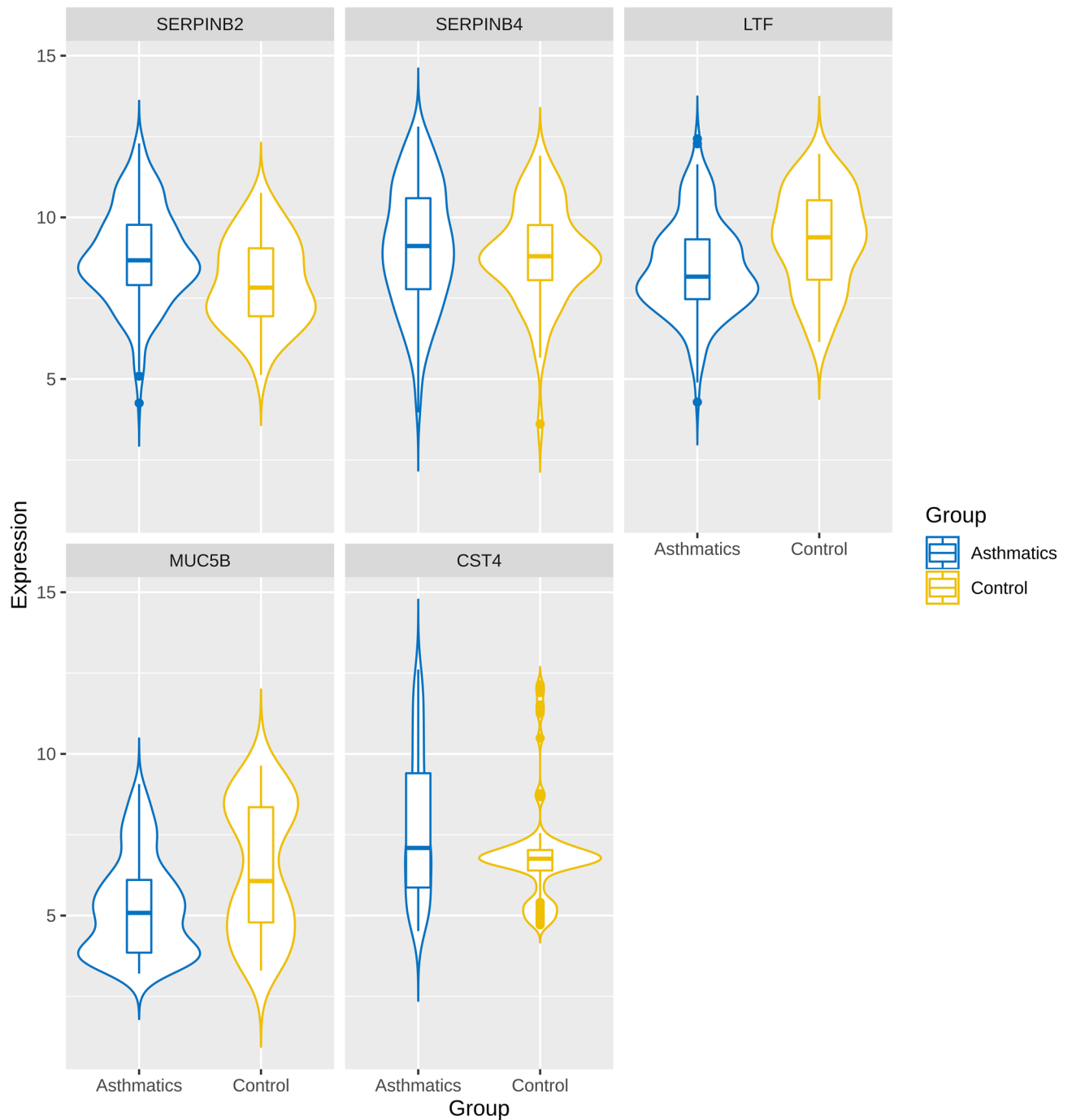


Figure 5. Violin plots of 5 hub genes filtered based on functional enrichment analysis and PPI network centrality measures respectively. This plot represents the distribution density of the underlying expression data for hub genes. The top and bottom of the embedded box signify the 75th and 25th percentile of the distribution, respectively. The line inside the box represents the median. Endpoints of the axis are labelled by the minimum and maximum values. Control and asthmatic samples are distinguished by yellow and blue colors, respectively.

studies^{35,36}. Thus targeting SERPINB4 as a localized molecular target leading to depletion of memory Th2 cells can help deplete the reservoir of inflammation in allergic asthma disease.

In asthma patients, TGF- β (Transforming growth factor-beta) increases connective tissue growth factor (CTGF) expression in airway smooth muscle cells which leads to deposition of extracellular matrix proteins, fibronectin, and collagen I³⁷. Wang et al.³⁸ revealed a TGF- β -SMAD4 and SMAD2 mediated epigenetic regulatory pathway for IL-9 production by differentiated Th9 cells. IL-9 is a proallergic cytokine which plays a key role in asthma induction. SMAD proteins may displace EZH2 directly from the IL9 locus by binding to it leading to IL-9 expression. The probability of cells to exhibit a sustained CTGF transcriptional response was increased by high total cellular SMAD4 but not SMAD2 levels³⁹. These studies pointed out the critical role of SMAD4 in

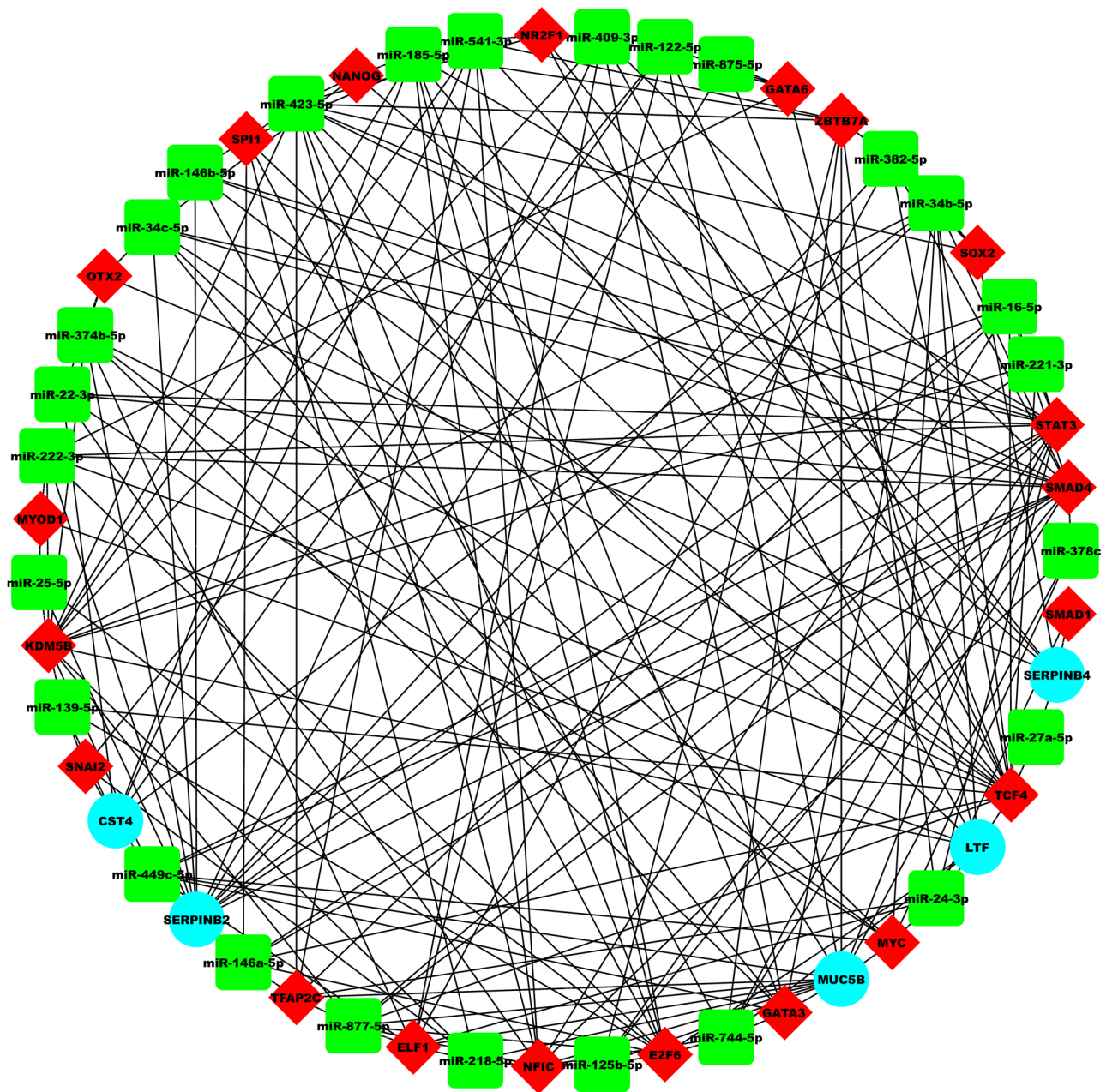


Figure 6. Asthma-specific 3-node miRNA FFL regulatory network with 51 nodes and 197 edges, respectively. The green-colored rectangular nodes represent asthma-associated miRNAs, red-colored triangular nodes represent transcription factors, and cyan-colored circular nodes represent asthma-associated hub genes.

Relationship	No. of edges	No. of miRNAs	No. of TFs	No. of genes
miRNA-gene ^a	32	26	-	5
TF-gene ^b	21	-	20	5
miRNA-TF ^c	144	26	20	-

Table 1. Summary of relationships among Asthma-associated DEGs, Asthma-associated miRNAs, and TFs. ^amiRNA-gene: miRNA repression of genes. ^bTF-gene: TF regulation of genes. ^cmiRNA-TF: miRNA repression of TFs.

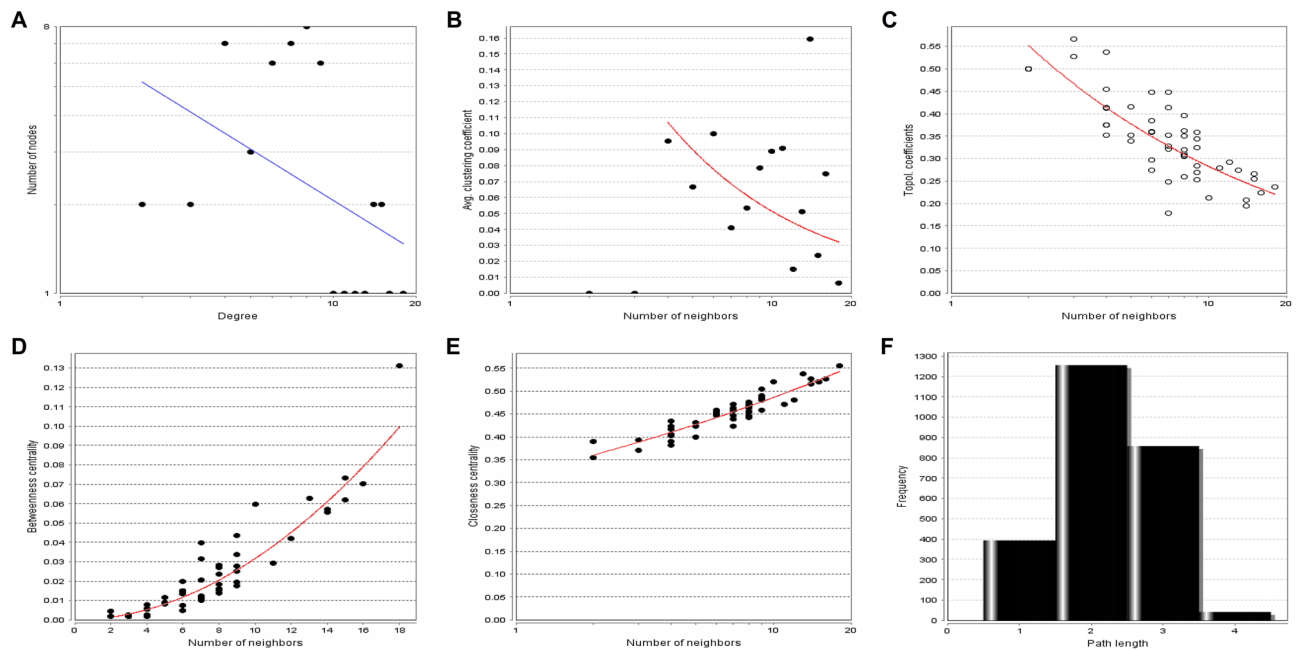


Figure 7. Topological parameter graphical plots of the asthma-specific 3-node miRNA FFL representing (A) Node degree distribution, (B) Average clustering coefficient, (C) Topological coefficient, (D) Betweenness centrality, (E) Closeness centrality, (F) Shortest path length distribution. The lines are fitted with power laws.

TGF- β mediated signaling in asthma⁴⁰. TGF- β -SMAD4 axis may be, therefore, explored as a novel therapeutic target to increase bronchodilators (mainly β_2 -adrenoceptor agonists) sensitivity in severe asthma. Our studies pointed out SMAD4 as a critical transcriptional factor involved in regulatory aspects of asthma pathogenesis. During fungal protease-mediated epithelial inflammation, downregulation of Uncoupling Protein 2 (UCP2) expression by regulatory TGF- β -SMAD4 signaling was associated with mitochondrial reactive oxygen species (ROS) production⁴¹.

Based on our analysis, we identified miR-34b-5p and miR-449c-5p (members of the miR-34/449 family) as highest-ranked microRNAs involved in the regulation of FFL involving transcription factor SMAD4 and SERPINB2 hub gene in asthma. Here based on the study it can be hypothesized that in asthma repression of miR 34b/449 regulates upregulation of SERPINB2 mediated via SMAD4 upregulation by an unknown direct or indirect regulatory process. Although this needs further validation to confirm the molecular mechanism underlying this. Although a few studies have been done to explore the role of miR-34b/449 in asthma but all of them indicate the key pathological role of miR 34b/449 in inflammation. An important study done by Solberg et al.⁴² shows that the miR-34/449 family members (miR-34c-5p, miR-34c-5p, miR-449a, and miR-449b-5p) are repressed significantly in vivo in asthma, in vitro by IL-13 exposure, and even after corticosteroid treatment. This repression of miRNA levels was induced by IL-13 resulting in increased notch expression as well as mucous cells increase in asthma. miR-449 acts as a crucial regulator of differentiation of nasal epithelial cells to ciliated cells via repression of transcript levels of NOTCH1 and its ligand Delta-1 (DLL1)^{42,43}. miR-34/449 also modulated Insulin-like growth factor-binding protein 3 (IGFBP-3) mediated autophagy activation in lung epithelial cells in vitro⁴⁴. Thus our study may provide an important link on miR-34b and miR-449 mediated regulation of Th2 cell signature gene SERPINB2 via IL-13. Given the fact that steroids have very little effect on miR expression in asthma, our study may help identify therapeutic targets which may lead to synchronized control in Th2-mediated inflammation. The study points out the possible role of 3',5'-cyclic adenosine monophosphate (cAMP) signaling pathways via SMAD4 as a critical regulator in asthma inflammation. This may be further mediated by increased PDE4 (Phosphodiesterase isomer) subsequently resulting in decreased Th2 biomarker like SERPINB2. miR-34b and miR-449 may be potential therapeutic targets which exert pathological inflammation by SMAD4 and SERPINB2. Exerting control by these two differentially expressed miRNAs can result in control in pathology exerted by airway epithelial cells in asthma.

The results obtained from our study exhibited the differential expression of hub genes along with their related mechanisms that could be used as potential therapeutic targets for asthma. Each hub gene and its respective partner may assist in gaining clear insights into the inflammation progress in mild, moderate, and severe asthma. Thereby, the differential expression pattern of hub genes reported in our study might help to gain insight into many unexplored regulatory pathways leading to unravelling the critical targets for increasing drug efficacy, diagnosis, and prognosis in asthma.

Methods

Data collection and preprocessing. Microarray datasets containing mRNA expression profiles for asthmatics and controls were downloaded from NCBI Genome Expression Omnibus (GEO)⁴⁵. The keyword “asthma” was used for this search. As per the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines⁴⁶, the eligible studies and datasets should follow these inclusion criteria: (i) having patient and control human samples, (ii) expression profiling by array type, (iii) having complete raw and processed microarray tissue data, (iv) submitted in year range of 2015 to 2020, and (v) derived from the same microarray platform. Review articles, abstracts, case reports, studies without healthy control or non-human samples, and cell-line based experimental study designs were excluded. GEOquery package⁴⁷ in R was used to download the CEL files which contained unprocessed mRNA expression profiles. Unprocessed expression data from CEL files were read as an affy-object using the affy module⁴⁸ followed by background correction, quantile normalization, and \log_2 transformation using Robust Multiarray Average (RMA) method in R. Also, the probe IDs were mapped to their corresponding HGNC (HUGO Gene Nomenclature Committee) gene symbol(s)⁴⁹ using hgu133plus2.db package available in R. Relative expression values across genes mapping to more than one probe IDs were averaged^{50,51}.

Meta-analysis of microarray datasets. The meta-analysis of the normalized gene expression datasets was performed in R using metaMA⁵² and limma⁵³ packages respectively. These packages employed functions for t-test and combined p value computation. The combined p value algorithm is based on Fisher’s sum of logs method⁵⁴ and sums the logarithm of p values across k -studies by combining datasets across samples for the same gene. The p values were adjusted using False Discovery Rate (FDR) method as given by Benjamini and Hochberg (BH)⁵⁵. Combined p values were computed for the common genes between both the datasets followed by their BH-correction. The DEGs between asthmatic and control samples were screened at BH - p -value < 0.05 and an absolute \log_2 (fold change) > 0.5 . The sets of significant genes corresponding to a \log_2 (fold change) > 0.5 and \log_2 (fold change) < -0.5 were categorized as the up and downregulated DEGs, respectively.

Pathway and GO term enrichment analysis. Pathway enrichment analysis was performed by different Enrichr^{56,57} gene set libraries like BioCarta, Reactome, Panther, NCI-Nature, and WikiPathways, respectively. To account for multiple comparisons, the pathways with BH - p -value < 0.05 were considered as significantly enriched. And, the GO term enrichment analysis was performed by Enrichr gene set libraries like GO-BP and GO-MF respectively. Terms with BH - p -value < 0.05 were considered as significantly enriched. Union of all DEGs involved in significantly enriched pathways and GO terms were used for further analysis.

PPI network analysis and hub gene identification. Unique DEGs obtained after enrichment analysis were used for PPI network construction and analysis. Interacting human protein partners of these DEGs were retrieved using the Biological General Repository for Interaction Datasets (BioGRID v3.5) (<https://thebiogrid.org/>)⁵⁸ and Human Integrated Protein–Protein Interaction rEference (HIPPIE v2.2) (<https://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/>)⁵⁹ databases respectively. The protein interactors having at least one reported experimental evidence were selected from BioGRID. And, the protein interactors with $0.63 \geq \text{score} \leq 1$ (corresponding to medium and high confidence interactions) were selected from HIPPIE. Both up and downregulated PPI networks were analyzed using the *NetworkAnalyzer* plugin available in Cytoscape v3.7.2⁶⁰. The *CytoHubba* plugin available in Cytoscape was used for hub genes identification⁶¹. Top n DEGs (where n is the number of DEGs in up and downregulated PPI networks) were ranked based on the five characteristic centrality measures like degree, stress, closeness, betweenness and EPC using *CytoHubba*. Intersecting DEGs from these five centralities for both the PPI networks were considered as the hub genes respectively.

Mining of significant interaction pairs between miRNAs, TFs, and hub genes. *miRNA-gene repression.* Retrieval of miRNA-gene pairs was performed using miRWalk v3.0⁶², miRSearch v3.0⁶³ and Starbase v2.0⁶⁴ databases respectively. Binding gap = 1, 3UTR region and score ≥ 0.95 were regarded as the threshold criteria for significant miRNAs selection in miRWalk respectively. Significant miRNAs having a high score denoted by the green-colored band were retrieved using miRSearch.

TF-gene regulation. TF-gene pairs retrieval was performed using ChIPBase v2.3 database⁶⁵, ITFP database⁶⁶, and Enrichr platform gene set libraries like ChEA, TRANSFAC, and JASPAR PWMs. TFs with binding sites location within 1 kb upstream region were extracted from ChIPBase. And, significant TFs with p value < 0.05 were extracted from Enrichr libraries.

miRNA-TF repression. The TFs obtained from TF-gene interaction were further utilized for fetching the relevant miRNA-TF pairs. TFs targeted by miRNAs were extracted by miRWalk, miRSearch, and Starbase databases respectively with the same thresholds described previously.

We required that the miRNA-target interactions be evolutionarily conserved in human and mouse. For this, mouse miRNAs targeting the available human TFs and genes were also retrieved using the same databases as described for miRNA-TF/gene interactions respectively. miRNAs common between human and mouse were considered as highly significant and validated (final miRNAs). All the three types of regulatory interactions were altered with respect to these final miRNAs. These molecular interactions were then merged to form a 3-node miRNA FFL⁶⁷ and subsequently visualized using Cytoscape.

Received: 3 February 2020; Accepted: 22 July 2020

Published online: 07 August 2020

References

- Enarson, D. *Respiratory Diseases in the World: Realities of Today—Opportunities for Tomorrow: Forum of International Respiratory Societies* (European Respiratory Society, Sheffield, 2013).
- Holgate, S. T. Pathogenesis of asthma. *Clin. Exp. Allergy J. Br. Soc. Allergy Clin. Immunol.* **38**(6), 872–897. <https://doi.org/10.1111/1/j.1365-2222.2008.02971.x> (2008).
- Brightling, C. E. *et al.* Mast-cell infiltration of airway smooth muscle in asthma. *N. Engl. J. Med.* **346**(22), 1699–1705. <https://doi.org/10.1056/NEJMoa012705> (2002).
- Munakata, M. Airway remodeling and airway smooth muscle in asthma. *Allergol. Int.* **55**(3), 235–243. <https://doi.org/10.2332/allergolint.55.235> (2006).
- James, A. L. *et al.* Airway smooth muscle hypertrophy and hyperplasia in asthma. *Am. J. Respir. Crit. Care Med.* **185**(10), 1058–1064. <https://doi.org/10.1164/rccm.201110-1849OC> (2012).
- Wenzel, S. E. Asthma: defining of the persistent adult phenotypes. *Lancet* **368**(9537), 804–813. [https://doi.org/10.1016/S0140-6736\(06\)69290-8](https://doi.org/10.1016/S0140-6736(06)69290-8) (2006).
- de Nijs, S. B., Venekamp, L. N. & Bel, E. H. Adult-onset asthma: is it really different?. *Eur. Respir. Rev.* **22**(127), 44–52. <https://doi.org/10.1183/09059180.00007112> (2013).
- Ray, A., Oriss, T. B. & Wenzel, S. E. Emerging molecular phenotypes of asthma. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **308**(2), L130–L140. <https://doi.org/10.1152/ajplung.00070.2014> (2015).
- Robinson, D. S. *et al.* Predominant TH2-like bronchoalveolar T-lymphocyte population in atopic asthma. *N. Engl. J. Med.* **326**(5), 298–304. <https://doi.org/10.1056/NEJM199201303260504> (1992).
- Wenzel, S. E. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat. Med.* **18**(5), 716–725. <https://doi.org/10.1038/nm.2678> (2012).
- Krug, N. *et al.* T-cell cytokine profile evaluated at the single cell level in BAL and blood in allergic asthma. *Am. J. Respir. Cell Mol. Biol.* **14**(4), 319–326. <https://doi.org/10.1165/ajrcmb.14.4.8600935> (1996).
- Kelly, E. A., Rodriguez, R. R., Busse, W. W. & Jarjour, N. N. The effect of segmental bronchoprovocation with allergen on airway lymphocyte function. *Am. J. Respir. Crit. Care Med.* **156**(5), 1421–1428. <https://doi.org/10.1164/ajrccm.156.5.9703054> (1997).
- Raman, K. Construction and analysis of protein-protein interaction networks. *Autom. Exp. 2*(1), 2. <https://doi.org/10.1186/1759-4499-2-2> (2010).
- Chen, S.-J., Liao, D.-L., Chen, C.-H., Wang, T.-Y. & Chen, K.-C. Construction and analysis of protein-protein interaction network of heroin use disorder. *Sci. Rep.* **9**(1), 4980. <https://doi.org/10.1038/s41598-019-41552-z> (2019).
- Estrada, E. Virtual identification of essential proteins within the protein interaction network of yeast. *Proteomics* **6**(1), 35–40. <https://doi.org/10.1002/pmic.200500209> (2006).
- Jeong, H., Mason, S. P., Barabási, A.-L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411**(6833), 41–42. <https://doi.org/10.1038/35075138> (2001).
- Shalgi, R., Brosh, R., Oren, M., Pilpel, Y. & Rotter, V. Coupling transcriptional and post-transcriptional miRNA regulation in the control of cell fate. *Aging* **1**(9), 762–770. <https://doi.org/10.18632/aging.100085> (2009).
- Shalgi, R., Lieber, D., Oren, M. & Pilpel, Y. Global and local architecture of the mammalian microRNA-transcription factor regulatory network. *PLoS Comput. Biol.* **3**(7), e131. <https://doi.org/10.1371/journal.pcbi.0030131> (2007).
- Harkema, J. R., Carey, S. A. & Wagner, J. G. The nose revisited: a brief review of the comparative structure, function, and toxicologic pathology of the nasal epithelium. *Toxicol. Pathol.* **34**(3), 252–269. <https://doi.org/10.1080/01926230600713475> (2006).
- Wu, Q., Qin, H., Zhao, Q. & He, X.-X. Emerging role of transcription factor-microRNA-target gene feed-forward loops in cancer. *Biomed. Rep.* **3**(5), 611–616. <https://doi.org/10.3892/br.2015.477> (2015).
- Liu, J. *et al.* Identification of hub genes and pathways associated with hepatocellular carcinoma based on network strategy. *Exp. Ther. Med.* **12**(4), 2109–2119. <https://doi.org/10.3892/etm.2016.3599> (2016).
- Poole, A. *et al.* Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. *J. Allergy Clin. Immunol.* **133**(3), 670–678.e12. <https://doi.org/10.1016/j.jaci.2013.11.025> (2014).
- Singhania, A. *et al.* Multitissue Transcriptomics Delineates the Diversity of Airway T Cell Functions in Asthma. *Am. J. Respir. Cell Mol. Biol.* **58**(2), 261–270. <https://doi.org/10.1165/rccm.2017-0162OC> (2018).
- Henskens, Y. M., Veerman, E. C. & Nieuw Amerongen, A. V. Cystatins in health and disease. *Biol. Chem. Hoppe. Seyler* **377**(2), 71–86. <https://doi.org/10.1515/bchm3.1996.377.2.71> (1996).
- Bonser, L. R. & Erle, D. J. Airway Mucus and Asthma: the role of MUC5AC and MUC5B. *J. Clin. Med.* **6**(12), 112. <https://doi.org/10.3390/jcm6120112> (2017).
- Fang, F. *et al.* Identification of potential transcriptomic markers in developing asthma: An integrative analysis of gene expression profiles. *Mol. Immunol.* **92**, 38–44. <https://doi.org/10.1016/j.molimm.2017.09.021> (2017).
- Nie, X. *et al.* Consistent biomarkers and related pathogenesis underlying asthma revealed by systems biology approach. *Int. J. Mol. Sci.* **20**(16), 4037. <https://doi.org/10.3390/ijms20164037> (2019).
- Fernández-Delgado, L. *et al.* Allergens induce the release of lactoferrin by neutrophils from asthmatic patients. *PLoS ONE* **10**(10), e0141278. <https://doi.org/10.1371/journal.pone.0141278> (2015).
- Kruzel, M. L., Bacsí, A., Choudhury, B., Sur, S. & Boldogh, I. Lactoferrin decreases pollen antigen-induced allergic airway inflammation in a murine model of asthma. *Immunology* **119**(2), 159–166. <https://doi.org/10.1111/j.1365-2567.2006.02417.x> (2006).
- Kucharewicz, I., Kowal, K., Buczek, W. & Bodzenta-Lukaszyk, A. The plasmin system in airway remodeling. *Thromb. Res.* **112**(1–2), 1–7. <https://doi.org/10.1016/j.thromres.2003.10.011> (2003).
- Woodruff, P. G. *et al.* T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am. J. Respir. Crit. Care Med.* **180**(5), 388–395. <https://doi.org/10.1164/rccm.200903-0392OC> (2009).
- Mertens, T. C. J., Hiemstra, P. S. & Taube, C. Azithromycin differentially affects the IL-13-induced expression profile in human bronchial epithelial cells. *Pulm. Pharmacol. Ther.* **39**, 14–20. <https://doi.org/10.1016/j.pupt.2016.05.005> (2016).
- Kuperman, D. A. *et al.* Direct effects of interleukin-13 on epithelial cells cause airway hyperreactivity and mucus overproduction in asthma. *Nat. Med.* **8**(8), 885–889. <https://doi.org/10.1038/nm734> (2002).
- Shamji, M. H. *et al.* Antiapoptotic serine protease inhibitors contribute to survival of allergenic TH2 cells. *J. Allergy Clin. Immunol.* **142**(2), 569–581.e5. <https://doi.org/10.1016/j.jaci.2017.07.055> (2018).
- Sivaprasad, U. *et al.* SERPINB3/B4 contributes to early inflammation and barrier dysfunction in an experimental murine model of atopic dermatitis. *J. Invest. Dermatol.* **135**(1), 160–169. <https://doi.org/10.1038/jid.2014.353> (2015).
- Ray, R. *et al.* Uteroglobin suppresses SCCA gene expression associated with allergic asthma. *J. Biol. Chem.* **280**(11), 9761–9764. <https://doi.org/10.1074/jbc.C400581200> (2005).
- Johnson, P. R. A. *et al.* Connective tissue growth factor induces extracellular matrix in asthmatic airway smooth muscle. *Am. J. Respir. Crit. Care Med.* **173**(1), 32–41. <https://doi.org/10.1164/rccm.200406-703OC> (2006).

38. Wang, A. *et al.* Cutting edge: Smad2 and Smad4 regulate TGF- β -mediated Il9 gene expression via EZH2 displacement. *J. Immunol.* **191**(10), 4908–4912. <https://doi.org/10.4049/jimmunol.1300433> (2013).
39. Tidin, O., Friman, E. T., Naef, F. & Suter, D. M. Quantitative relationships between SMAD dynamics and target gene activation kinetics in single live cells. *Sci. Rep.* **9**(1), 5372. <https://doi.org/10.1038/s41598-019-41870-2> (2019).
40. Wortley, M. A. & Bonvini, S. J. Transforming growth factor- β 1: a novel cause of resistance to bronchodilators in Asthma?. *Am. J. Respir. Cell Mol. Biol.* **61**(2), 134–135. <https://doi.org/10.1165/rcmb.2019-0020ED> (2019).
41. Kim, Y. H. & Lee, S.-H. TGF- β /SMAD4 mediated UCP2 downregulation contributes to Aspergillus protease-induced inflammation in primary bronchial epithelial cells. *Redox Biol.* **18**, 104–113. <https://doi.org/10.1016/j.redox.2018.06.011> (2018).
42. Solberg, O. D. *et al.* Airway epithelial miRNA expression is altered in asthma. *Am. J. Respir. Crit. Care Med.* **186**(10), 965–974. <https://doi.org/10.1164/rccm.201201-0027OC> (2012).
43. Marcet, B. *et al.* Control of vertebrate multiciliogenesis by miR-449 through direct repression of the Delta/Notch pathway. *Nat. Cell Biol.* **13**(6), 693–699. <https://doi.org/10.1038/ncb2241> (2011).
44. Yin, H. *et al.* MicroRNA-34/449 targets IGFBP-3 and attenuates airway remodeling by suppressing Nur77-mediated autophagy. *Cell Death Dis.* **8**(8), e2998. <https://doi.org/10.1038/cddis.2017.357> (2017).
45. Clough, E. & Barrett, T. The Gene Expression Omnibus Database. *Methods Mol. Biol. Clifton NJ* **1418**, 93–110. https://doi.org/10.1007/978-1-4939-3578-9_5 (2016).
46. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & for the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* **339**, b2535–b2535. <https://doi.org/10.1136/bmj.b2535> (2009).
47. Davis, S. & Meltzer, P. S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinforma. Oxf. Engl.* **23**(14), 1846–1847. <https://doi.org/10.1093/bioinformatics/btm254> (2007).
48. Irizarry, R. A. *et al.* Summaries of Affymetrix GeneChip probe level data. *Nucl. Acids Res.* **31**(4), e15. <https://doi.org/10.1093/nar/gng015> (2003).
49. Gray, K. A., Seal, R. L., Tweedie, S., Wright, M. W. & Bruford, E. A. A review of the new HGNC gene family resource. *Hum. Genomics* **10**, 6. <https://doi.org/10.1186/s40246-016-0062-6> (2016).
50. Ahmad, S. *et al.* Transcriptome meta-analysis deciphers a dysregulation in immune response-associated gene signatures during sepsis. *Genes* **10**(12), 1005. <https://doi.org/10.3390/genes10121005> (2019).
51. Shriwasth, N. *et al.* Identification of differentially expressed genes in small and non-small cell lung cancer based on meta-analysis of mRNA. *Heliyon* **5**(6), e01707. <https://doi.org/10.1016/j.heliyon.2019.e01707> (2019).
52. Marot, G., Foulley, J.-L., Mayer, C.-D. & Jaffrézic, F. Moderated effect size and P-value combinations for microarray meta-analyses. *Bioinforma. Oxf. Engl.* **25**(20), 2692–2699. <https://doi.org/10.1093/bioinformatics/btp444> (2009).
53. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-seq and microarray studies. *Nucl. Acids Res.* **43**(7), e47–e47. <https://doi.org/10.1093/nar/gkv007> (2015).
54. Fisher, R. A. Statistical Methods for Research Workers. In *Breakthroughs in Statistics* (eds Kotz, S. & Johnson, N. L.) 66–70 (Springer, New York, NY, 1992).
55. Benjamini, Y. Discovering the false discovery rate: False Discovery Rate. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **72**(4), 405–416. <https://doi.org/10.1111/j.1467-9868.2010.00746.x> (2010).
56. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14**, 128. <https://doi.org/10.1186/1471-2105-14-128> (2013).
57. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucl. Acids Res.* **44**(W1), W90–97. <https://doi.org/10.1093/nar/gkw377> (2016).
58. Oughtred, R. *et al.* The BioGRID interaction database: 2019 update. *Nucl. Acids Res.* **47**(D1), D529–D541. <https://doi.org/10.1093/nar/gky1079> (2019).
59. Alanis-Lobato, G., Andrade-Navarro, M. A. & Schaefer, M. H. HIPPIE v2.0: enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucl. Acids Res.* **45**(D1), D408–D414. <https://doi.org/10.1093/nar/gkw985> (2017).
60. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**(11), 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
61. Sriroopreddy, R., Sajeed, R., Raghuraman, P., Sudandiradoss, C. Differentially expressed gene (DEG) based protein-protein interaction (PPI) network identifies a spectrum of gene interactome, transcriptome and correlated miRNA in nondisjunction Down syndrome. *Int. J. Biol. Macromol.* **122**, 1080–1089. <https://doi.org/10.1016/j.ijbiomac.2018.09.056> (2019).
62. Sticht, C., De La Torre, C., Parveen, A. & Gretz, N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS ONE* **13**(10), e0206239. <https://doi.org/10.1371/journal.pone.0206239> (2018).
63. Açıcı, K., Terzi, Y. K. & Oğul, H. Retrieving relevant experiments: The case of microRNA microarrays. *Biosystems* **134**, 71–78. <https://doi.org/10.1016/j.biosystems.2015.06.003> (2015).
64. Li, J.-H., Liu, S., Zhou, H., Qu, L.-H. & Yang, J.-H. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucl. Acids Res.* **42**(D1), D92–D97. <https://doi.org/10.1093/nar/gkt1248> (2014).
65. Zhou, K.-R. *et al.* CHIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. *Nucl. Acids Res.* **45**(D1), D43–D50. <https://doi.org/10.1093/nar/gkw965> (2017).
66. Zheng, G. *et al.* ITPP: an integrated platform of mammalian transcription factors. *Bioinforma. Oxf. Engl.* **24**(20), 2416–2417. <https://doi.org/10.1093/bioinformatics/btn439> (2008).
67. Sun, J., Gong, X., Purow, B. & Zhao, Z. Uncovering microRNA and transcription factor mediated regulatory networks in glioblastoma. *PLoS Comput. Biol.* **8**(7), e1002488. <https://doi.org/10.1371/journal.pcbi.1002488> (2012).

Acknowledgements

We would like to thank Jamia Millia Islamia for providing infrastructure, journal access, and internet facilities. We acknowledge the SERB (Science and Engineering Research Board), DST (Department of Science and Technology, India) for providing the Ramanujan Fellowship (Grant ID: SB/S2/RJN-199/2014) to Mansoor Ali Syed and financial assistance (Grant ID: EEQ/2016/000509) to Ravins Dohare. Archana Sharma and Shweta Arora would like to thank ICMR (Indian Council of Medical Research) for providing Research Associate Fellowship and Senior Research Fellowship, respectively.

Author contributions

M.A.S. and R.D. conceived the study. P.S., A.S., and R.J. analysed the data and interpreted the results. A.S., S.A., P.S., and R.J. wrote the manuscript. P.S. and R.J. prepared all the figures. R.A., A.H.R., and S.A.A. reviewed and edited the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-70368-5>.

Correspondence and requests for materials should be addressed to R.D. or M.A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020