Integrating Genetics, Transcriptomics, and Proteomics in Lung Tissue to Investigate COPD

Yu-Hang Zhang[1], Michael H. Cho[1,2], Peter J. Castaldi[1], Craig P. Hersh[1,2], Mukul K. Midha[3], Michael R. Hoopmann[3], Robert L. Moritz[3], Edwin K. Silverman[1,2]

[1] Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA ; [2] Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA ; [3] Institute for Systems Biology, Seattle, WA, USA

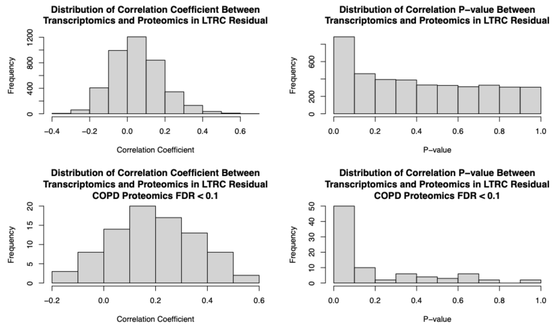
Rationale

Omics data at multiple biological levels can enable comprehensive analyses of biological processes. We integrated lung tissue **transcriptomics** and **proteomics** data with **genetic variation** to provide new insights into COPD pathogenesis.

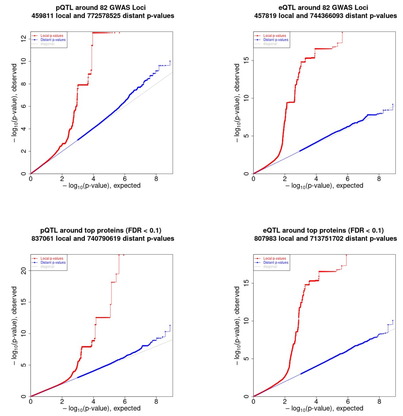
Methods

i. Lung tissue samples from 98 Lung Tissue Research Consortium subject were analyzed with whole genome sequencing, RNA-seq, and mass spectrometry proteomics
ii. Pearson's correlation test for Omics expression comparison
iii. QTL analyses on genomic regions of interest
iv. Mediation analyses linking COPD GWAS loci, COPD biomarker gene expression and COPD
v. Colocalization analyses on GWAS, eQTL and pQTL signals
vi. WGCNA preservation analyses

1. Low correlations between transcriptomics and proteomics were observed overall ,but COPD-associated proteins had stronger correlations.



2. QQ plots for QTL analyses



3. Significant QTL analyses results

Significant cis-QTL effects were observed near multiple COPD biomarkers including eleven cis eQTLs and five pQTLs with FDR threshold as 0.1.

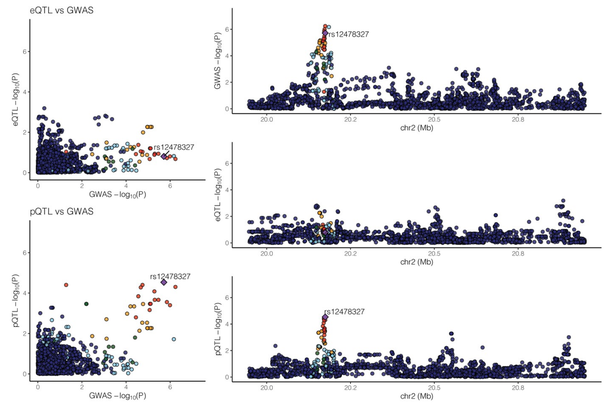
Results

4. Significant mediation effects were found with P-value threshold as 0.1

GWAS SNP	Gene Name	UniProt ID	Omics Level	Mediation Proportion	Mediated effects P	Mediated effects beta (se)	Total effects P	Total effects beta (se)	QTL analyses P	QTL analyses beta (se)
rs8080772	*AQP4*	P55087	transcriptomics	74% (-203%,473%)	4.10E-02	0.084(0.041)	8.52E-02	0.114(0.066)	5.23E-02	-0.181(0.092)
rs8080772	*VAMP3*	Q15836	proteomics	66% (-184%,446%)	3.01E-02	0.074(0.034)	7.53E-02	0.112(0.063)	1.38E-03	-0.382(0.116)
rs9435731	*AQP1*	P29972	proteomics	60% (-191%,391%)	4.52E-02	-0.063(0.031)	9.44E-02	-0.105(0.063)	6.01E-02	0.191(0.101)
rs8080772	*HSPB1*	P04792	proteomics	53% (-111%,308%)	4.22E-02	0.057(0.028)	8.28E-02	0.109(0.063)	3.83E-02	0.134(0.064)

5. Colocalization analyses

A suggestive colocalization with common causal variants between the pQTL and COPD GWAS signal near RHOB was identified with colocalized probability of 0.562.



6. Weighted Gene Correlation Network Analysis (WGCNA) preservation

Two preserved network modules (gene clusters) generated by WGCNA were associated with COPD (FDR < 0.05). One module is related to the catenin complex and the other module to plasma membrane components.

Conclusion

We evaluated associations between lung tissue Omics. Despite insufficient power to detect long-range genetic effects, multiple cis-acting QTL effects were identified. Mediation analysis, and correlation-based network analysis of multiple Omics data may identify key genes and proteins that influence COPD pathogenesis.

Acknowledgement

We thank the study participants in the Lung Tissue Research Consortium