Identifying the gene co-expression module for tobacco smoking that is regulated by DNA methylation

Ke Xu[1, 2], Xinyu Zhang[1,2], Ying Hu[4], Hongyu Zhao[5], Rajita Sinha[1,3]

[1]Department of Psychiatry, Yale School of Medicine, 300 George street, suite 901, New Haven, CT 06511; [2]Connecticut Veteran Health System, 950 Campbell Ave, Building 35, Room #43, West Haven, 06516; [3]Yale Stress Center, Yale University, 2 Church St S #209, New Haven, CT 06519; [4]Center for Biomedical Informatics and Information Technology, National Cancer Institute, 9609 Medical Center Drive, Rockville, Maryland 20850; [5]Department of Biostatistics, Yale School of Public Health, New Haven, CT, 06511

Background: Tobacco smoking has significant impacts on epigenome and transcriptome across different tissues and cell types. Previous transcriptomic studies have identified individual differentially expressed genes between smokers and non-smokers. However, the relationship of these genes at the system level has not been well studied. Whether differential gene expression for smoking is modulated by epigenetic modification is still unknown. **Method**: We profiled transcriptome expression (Illumina HumanHT-12 v4 Expression BeadChip) and genome-wide DNA methylation (Illumina HumanMethylation 450K BeadChip) in blood samples of 506 healthy individuals with and without smoking. A weighted gene co-expression network analysis (WGCA), an unsupervised network analysis approach, was applied to identify gene expression clusters and to test the relationship between co-expressed genes and smoking. We then examined correlation between DNA methylation and gene expression in the smoking-associated gene co-expressed module. **Results:** Using WGCNA, we identified 25 co-expressed gene modules from top 5,000 variable genes in all 506 samples. One co-expression gene module (Salmon) containing 88 genes was significantly associated with smoking ($p = 0.013$). Four out of 88 genes differentially expressed between smokers and non-smokers (corrected p values ~ 0.008 – 0.002). Two hub genes in this module, *RNF13* (ring finger protein 13) and *GCA* (EF-hand calcium-binding protein) were highly connected with other genes. RNF13 is a critical mediator that facilitates stress-induced apoptosis by activating the IRE1α-TRAF2-JNK signaling pathway. The cluster of genes connected to GCA were in the immune function pathway. Genes in the smoking-associated module were significantly enriched in multiple pathways related to the adaptive immune response, regulation of defense, and regulation of stress responses by a hypergeometric test (FDR $q < 0.05$). More importantly, DNA methylation on 8 genes in this module was significantly correlated with gene expression ($p < 0.05$). Particularly, methylation of the hub gene, *RNF13*, was inversely correlated with expression, indicating that DNA methylation regulates gene co-expression of this smoking-associated module. **Conclusion**: Our results suggest that smoking-altered transcriptome acts as a functional group that may be modulated by DNA methylation. Our findings provide molecular insights of smoking in healthy individuals and implicate a significant role for epigenetic regulation of gene function in nicotine-related pathologies.