Submitter Name: Fang Chen
Submitted email: fchen1@hmc.psu.edu
PI Name (if different): Dajiang Liu
PI email (if different): dxl46@psu.edu

Trans-ethnic Transcriptome-wide Association Study for Smoking Addiction in 1.3 Million Individuals Yields Insights into Tobacco Use Biology and Drug Repurposing

Fang Chen[1], Xingyan Wang[1], Seon-Kyeong Jang[2], J. Dylan Weissenkampen[1], Chachrit Khunsriraksakul[1], Dana Hancock[3], Bibo Jiang[1], Scott Vrieze[2], Dajiang J. Liu[1]

[1]Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, USA;[2] Department of Psychology, University of Minnesota, Minneapolis, MN, USA;[3]RTI International, USA

Genome-wide association studies using samples of European ancestry have discovered more than 400 loci associated with tobacco use behaviors, with a majority of the associated variants being non-coding. It remains challenging to map these non-coding variants to their target genes and translate their biological and clinical relevance. Transcriptome-wide association studies (TWAS) address this issue by integrating genetically regulated gene-expression data from ancestry-matched eQTL datasets. As genetic studies start to incorporate samples of non-European studies, it is critical to extend TWAS accordingly for trans-ethnic genetic studies.

Here we aggregated GWAS and whole genome sequencing data from TOPMed (total N = 1.3 million) and eQTL data in multiple tissues from diverse ancestries to further empower gene discovery for tobacco use behaviors. We developed a novel approach, TESLA, that optimally integrates trans-ethnic GWAS with eQTL datasets. TESLA greatly outperforms prior TWAS methods that integrate only GWAS and eQTL data with matched ancestry, and that incorporate fixed effect GWAS meta-analysis results. The advantage of TESLA is consistent regardless of the gene expression prediction models used. Applying TESLA to tobacco use phenotypes, we identified 319 novel genes that are outside 1 million base pair windows of GWAS sentinel variants, and suggested key pathways that are ubiquitous across tissues (e.g., neurotransmitter catabolic process, GABAergic and dopaminergic synapse) and pathways that are more tissue-specific. Computational drug repurposing using TESLA results also highlighted several drugs with known efficacy including dextromethorphan and galantamine, and novel drugs such as muscle relaxant that may be repurposed for treating smoking addiction.