Name: Madhurbain Singh
Email: singhm18@vcu.edu
PI Name: Michael C. Neale
PI email: michael.neale@vcuhealth.edu

Examining Bidirectional Causal Effects between Smoking and DNA Methylation using Epigenetic Mendelian Randomization Analyses

Madhurbain Singh[1,2], Conor V. Dolan[3,4], Jouke Jan Hottenga[3], Rene Pool[3], Hermine H. Maes[1,2], Brad Verhulst[5], Gibran Hemani[6], Josine L. Min[6], BIOS Consortium[7], Eco J. C. de Geus[3,4], Dorret I. Boomsma[3,4], Jenny van Dongen[3,4,9], and Michael C. Neale[1,2,3,8,9]

[1]Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA; Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA; [2]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [3]Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; [4]Department of Psychiatry and Behavioral Sciences, Texas A&M University, College Station, TX; [5]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK; [6]Biobank-based Integrative Omics Study Consortium; [7]Department of Psychiatry, Virginia Commonwealth University, Richmond, VA; [8]These authors jointly supervised this work.

Prior epigenome-wide association studies have found widespread robust associations between cigarette smoking and DNA methylation (DNAm) in peripheral blood cells. The etiology of these associations plausibly varies across CpG sites, arising from 1) the causal effects of smoking on DNAm, 2) the impact of DNAm on smoking liability, or 3) genetic and environmental confounders influencing both smoking and DNAm. To disaggregate potential bidirectional causal effects between smoking and DNAm, we combined Mendelian Randomization analyses with twin data from the Netherlands Twin Register (N = 514 current, 540 former, and 1424 never smokers of European ancestry). We focused on ~18,000 smoking-related CpG sites previously identified. We also conducted sensitivity analyses, adjusting for horizontal pleiotropy and the residual genetic and environmental correlations. Triangulating across models, we find a significant effect of DNAm levels on smoking liability at >40 CpG sites (after multiple-testing correction), while current (vs. never) smoking has a significant effect on DNAm at only eight sites. Moreover, the estimated effect of DNAm levels on smoking liability is usually stronger than the reverse effect of smoking on DNAm. The effect of former (vs. never) smoking on DNAm is further attenuated across CpGs, consistent with reversibility of smoking's effects. We find bidirectional effects between smoking and DNAm in/near the *AHRR* and *ALPPL2* genes, indicating possible feedback mechanisms. Beyond these genes, there is limited overlap between the CpG sites with unidirectional causal effects, thus differentiating between potential biomarkers of current or past smoking exposure and biomarkers of smoking liability.