Results from the Decode GWAS study on addiction

Thorgeir E. Thorgeirsson, Gunnar W. Reginsson[1], Gyda Bjornsdottir[1], Thorarinn Tyrfingsson[2], Valgerdur Runarsdottir[2], Ingunn Hansdottir[2,3], Stacy Steinberg[1], Hreinn Stefansson[1], Daniel F. Gudbjartsson[1,4], Kari Stefansson[1,5]

[1]deCODE genetics / Amgen, Sturlugata 8, Reykjavik, Iceland, [2]SAA-National Center of Addiction Medicine, Reykjavik, Iceland, [3]Faculty of Psychology, University of Iceland, Reykjavik, Iceland, [4]Department of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland, [5]Faculty of Medicine, University of Iceland, Reykjavik, Iceland

The identification of the first genome-wide significant variant correlating with CPD required 10,000 samples, with only a few more variants identified in studies of ~85,000samples, and recently as part of the GSCAN study, a sample size of ~337,000 yielded 40 variants correlating with CPD. A recent GWAS of alcohol consumption in UK biobank data found 8 loci in ~112,000 samples, and GSCAN identified 68 variants at a sample size near one million. The yield from GWASs of smoking initiation was only a few hits at ~100,000 cases and controls, but GSCAN identified 262 variants in 1.2 million cases and controls. Gene identification in Schizophrenia required smaller sample size, with ~130 variants identified using ~37,000 cases. It is not unlikely that the sample size required for alcohol and drug addiction will fall somewhere between that required for SCZ and those required for substance use phenotypes. Given the correlation between addiction and the polygenic risk of SCZ, a substantial fraction of the variants to be identified can be expected to also correlate with SCZ, and some perhaps with other psychiatric disorders.

We have performed a GWAS of addictive disorders in Iceland comparing allele frequencies for over 50 million variants in 21,339 subjects treated for addiction at the SAA treatment center between the years 1977 and 2014 to those in 294,099 population controls. The analysis is based on genotype data containing 15,522 whole genome sequences (WGS) (at over 30x depth) directly imputed into 151,677 samples with accurate haplotypes based on long-range phasing, as well as genotypes based on familial imputation for 282,894 additional subjects.

Two genome-wide significant loci were identified. Neither locus seems very substance specific, as common variants at both loci confer risk of several substance use disorders, and one of the key variants also confers risk of depression, with a similar effect size for addiction and depression. Associations with consequences of excessive alcohol use on health were also discovered, highlighting the public health relevance of addiction risk variants. The WGS-imputation approach allows the study of a large number of rare variants with some power to detect associations. These results will be discussed. For example, we have identified a truncating mutation in *RBM12* associating strongly with both psychosis and addiction, with a ratio of effect sizes similar to that observed for the entire SCZ polygenic score.

This study was funded in part by the National Institute on Drug Abuse (NIDA) (R01-DA017932 and R01-DA034076)