SUPPLEMENTARY MATERIAL

Genome-wide association study of suicide death and polygenic prediction of clinical antecedents

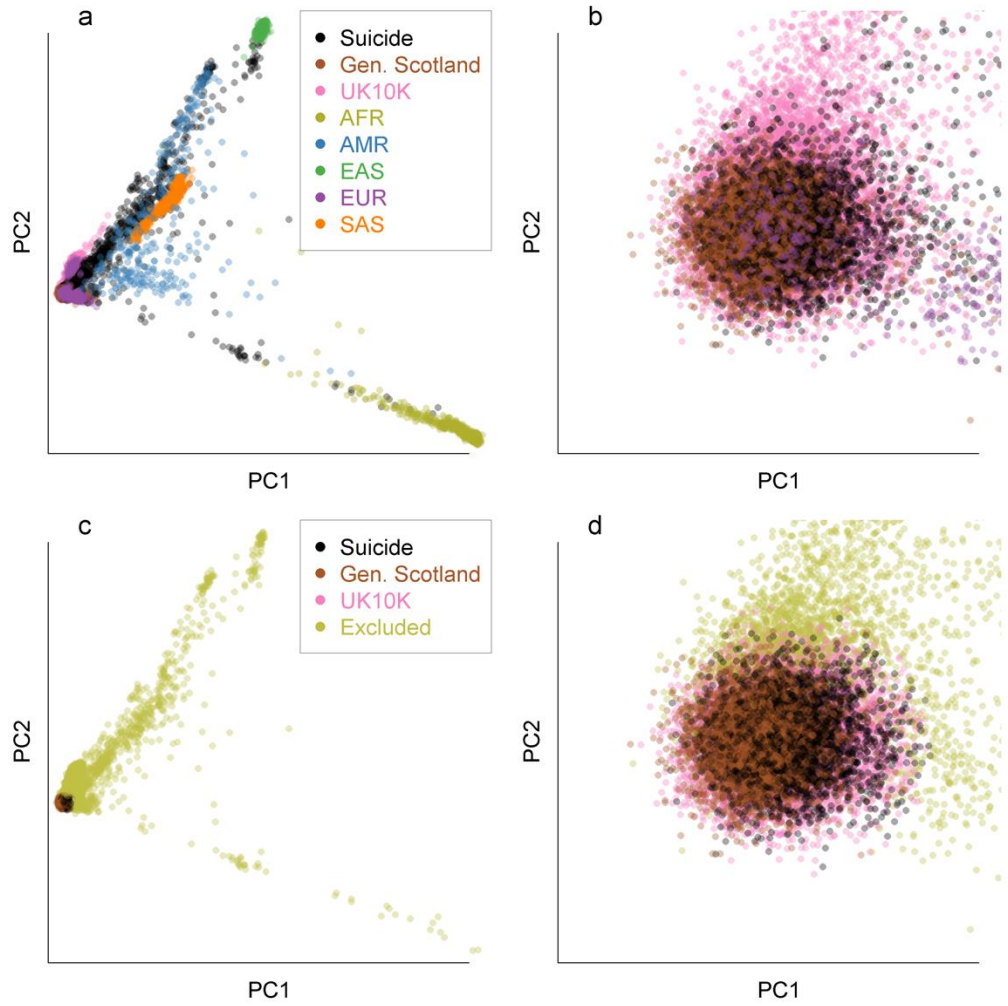
SUPPLEMENTARY METHODS

S1. Additional Details of Sample Collection & Genotyping. In the first wave we genotyped 1,321 cases, including samples from the beginning of the collection through January 2014. Funding limitations precluded genotyping of all samples. This wave 1 genotyping prioritized cases with the most complete available data from the Utah Population Database, including completeness of diagnostic and genealogical data. Subsequent collection and funding resources made wave 2 possible, and this genotyping included the samples not previously genotyped in wave 1, along with the additional, newer cases made available 2014-2017.

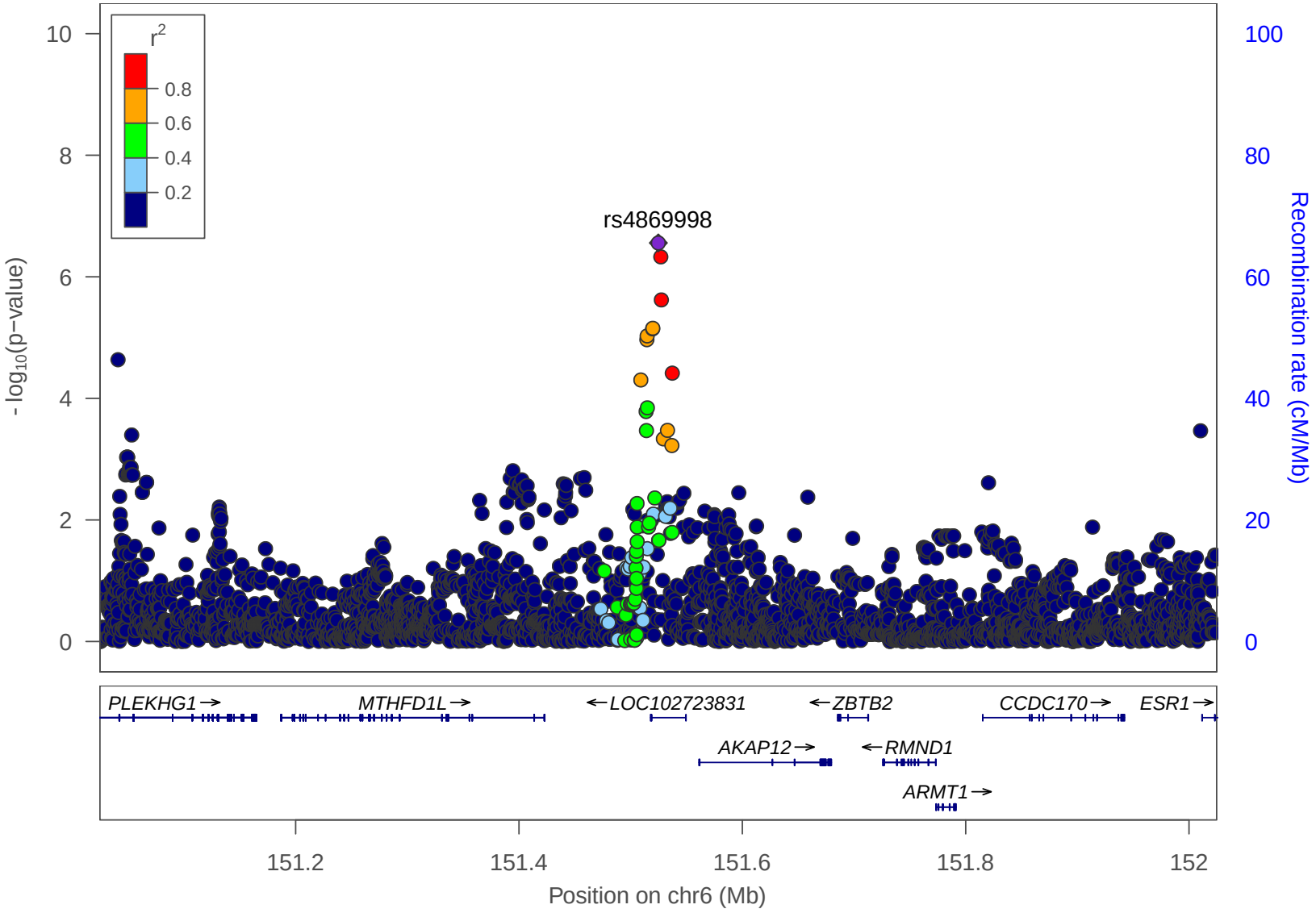
S2. International Classification of Diseases (ICD-10) Code Ascertainment. Decedent data are linked within the Utah Population Database (UPDB), a unique resource that houses data on > 9 million individuals and contains vital statistic and demographic records and electronic health records from the two major hospital systems in the Utah.[22] ICD codes for suicide cases were recovered from the electronic data warehouses of Utah's two largest health care providers, Intermountain Healthcare and the University of Utah. Ambulatory and inpatient ICD codes were obtained directly from UPDB. EMR data, while comprehensive and population-based, are also subject to missingness (random and non-random), and severity of health disparity may correspond with the presence and absence of ICD codes in an EMR. For this reason, we checked ICD code associations with polygenic risk for the same disorders and assessed relationships of ICD with PRS across age group and date of code. The rationale for the latter analyses was that younger individuals are less likely to have specific ICD codes due to either lack of contact with the system or to lack of maturation of overt psychopathology (e.g., psychosis, personality disorders). In addition, ICD codes from early EMR development in Utah (pre-year 2000) are notably sparser. To manage the wide diversity of codes, we limited our analyses to 30 categories of codes reflecting 30 relevant diagnostic domains. Data were subsequently completely de-identified prior to analysis. The study was approved by the Institutional Review Boards of the University of Utah, Intermountain Healthcare, and the Utah Department of Health.

S3. Clinical Diagnostic Associations with Mode of Death are presented in Supplementary Figure S9. Suicide by gun was associated with the general absence of clinical diagnoses. Suicide by overdose was associated with many clinical diagnoses and most robustly with obesity and sleep disorders. Suicide by violent trauma was associated with a clinical diagnosis of schizophrenia. Corresponding βs and p-values for Figure 5 are presented in Supplementary Tables S11-S13.

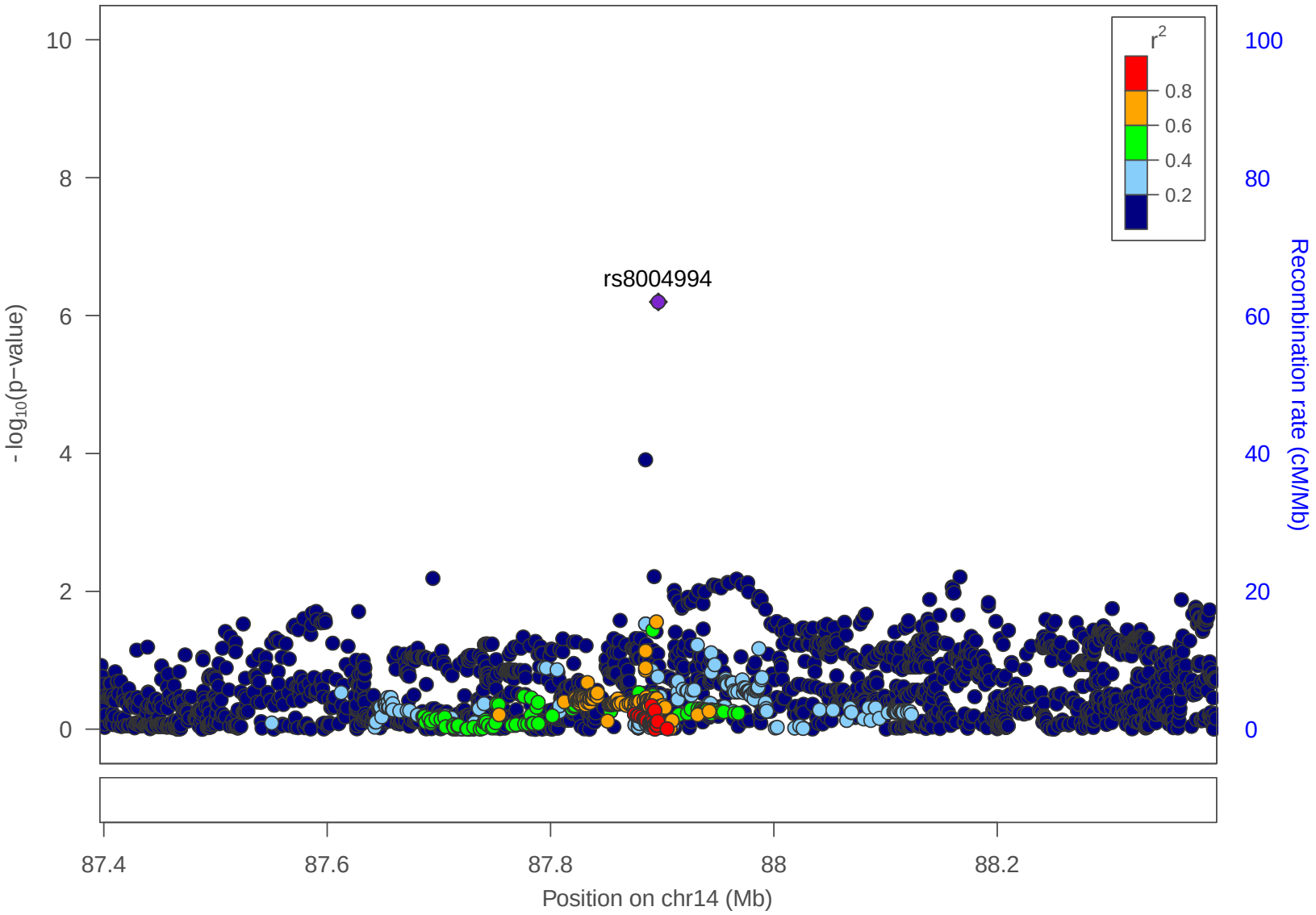
SUPPLEMENTARY FIGURES



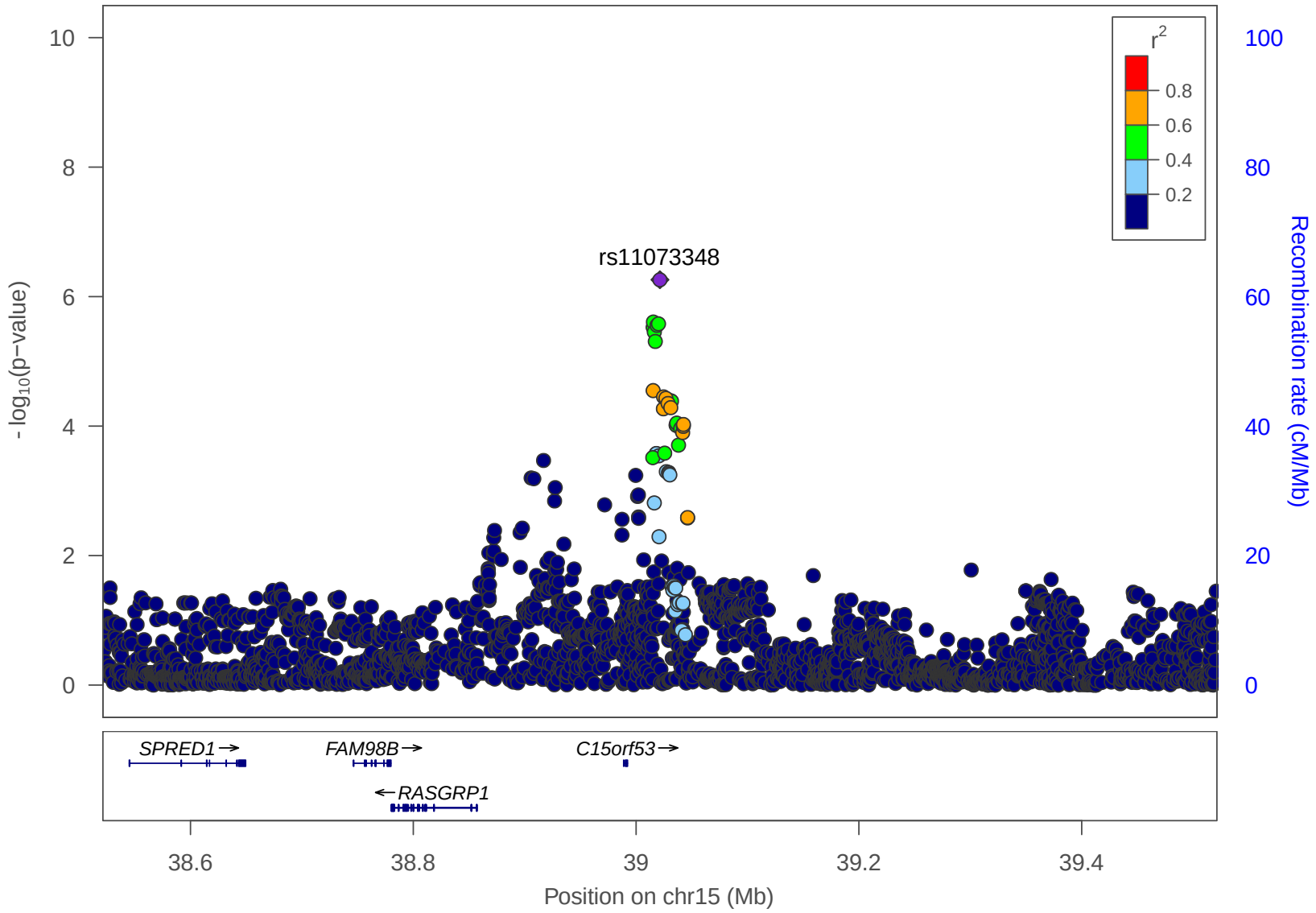
S1. PCA of 1KG super-populations, cases, controls, and excluded samples. Suicide death samples plotted by the principal component explaining the most variance (PC1) versus the principal component explaining the second most variance (PC2) in all 1KG super-populations (a) and focusing on only the Northern European cases and controls (b). (c) and (d) highlight excluded cases. For adequate statistical power, we examined only cases of Northern European ancestry. However, it is clear from (a) and (c) that the cohort was comprised of multiple ancestries and that research on suicide death in non-European ancestries will reflect an important step beyond this first study.



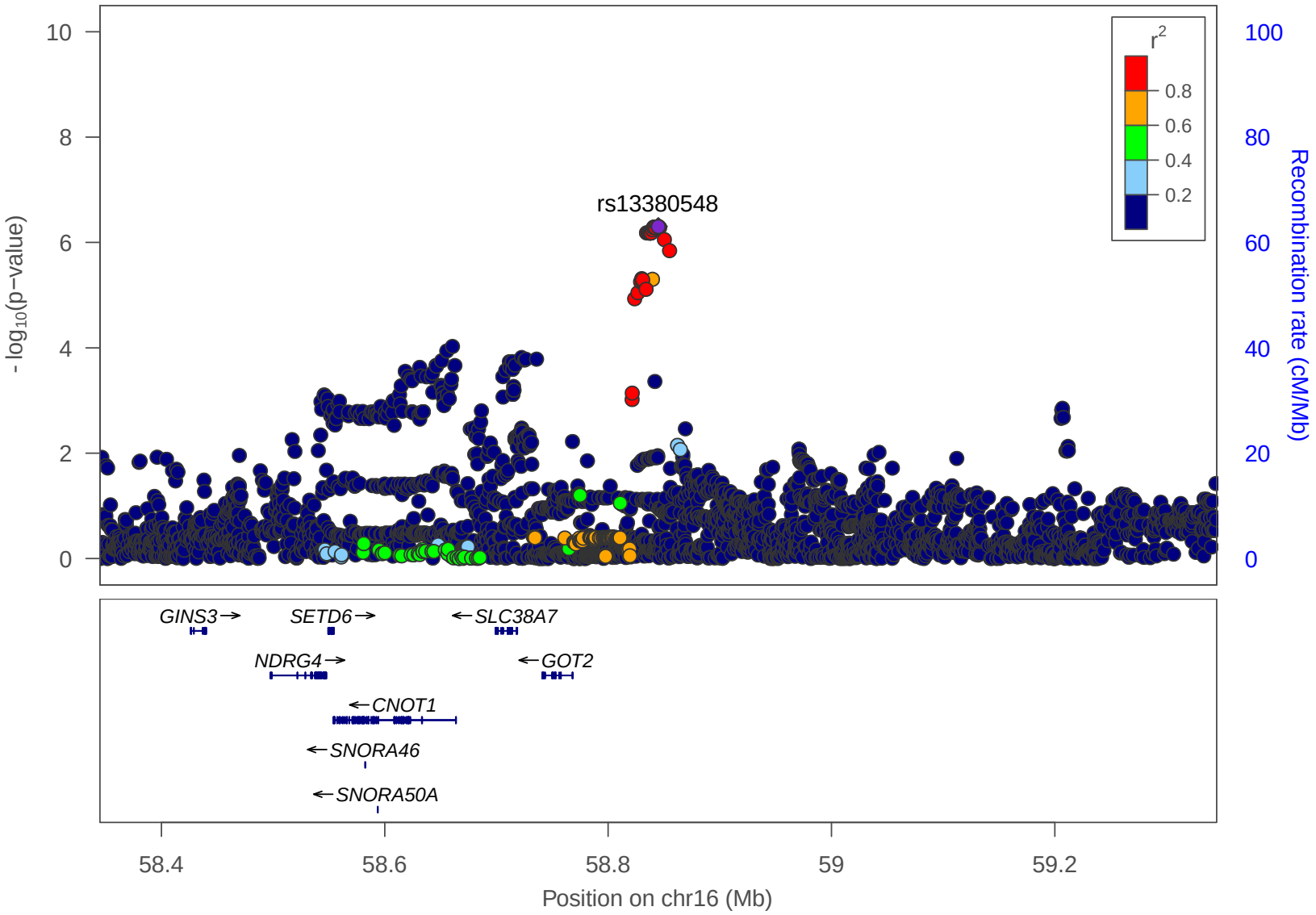
S2. Chromosome 6



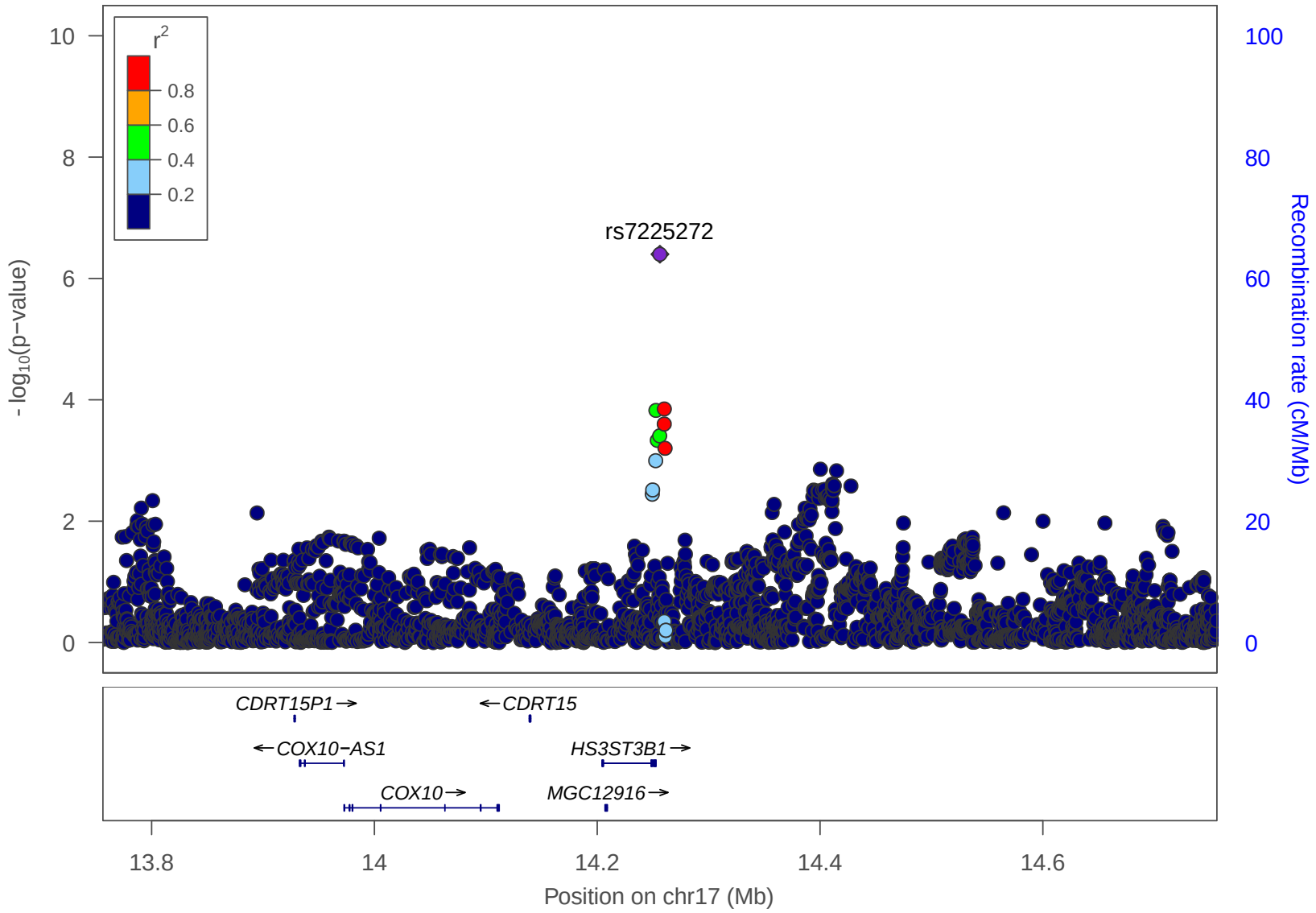
S3. Chromosome 14



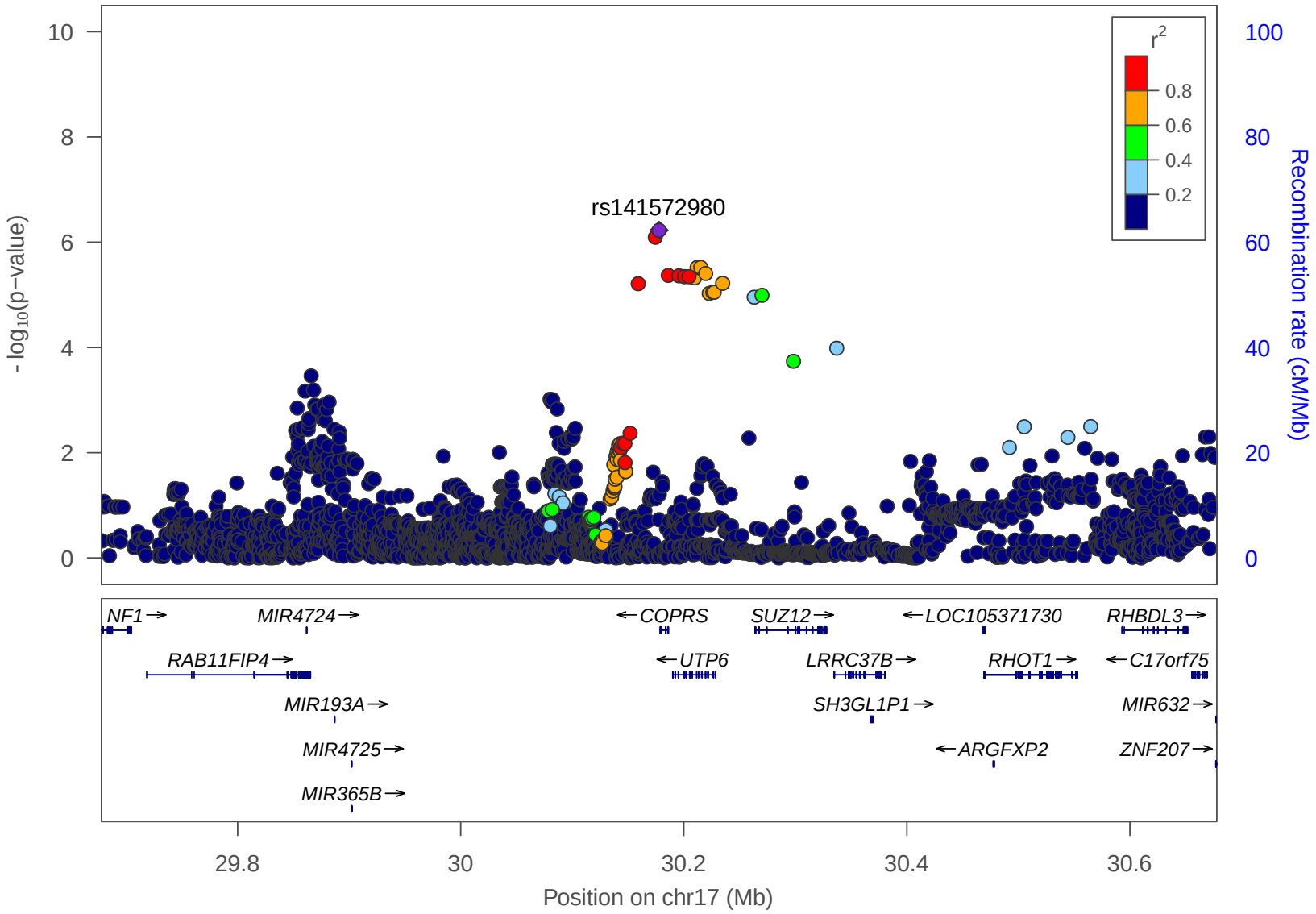
S4. Chromosome 15



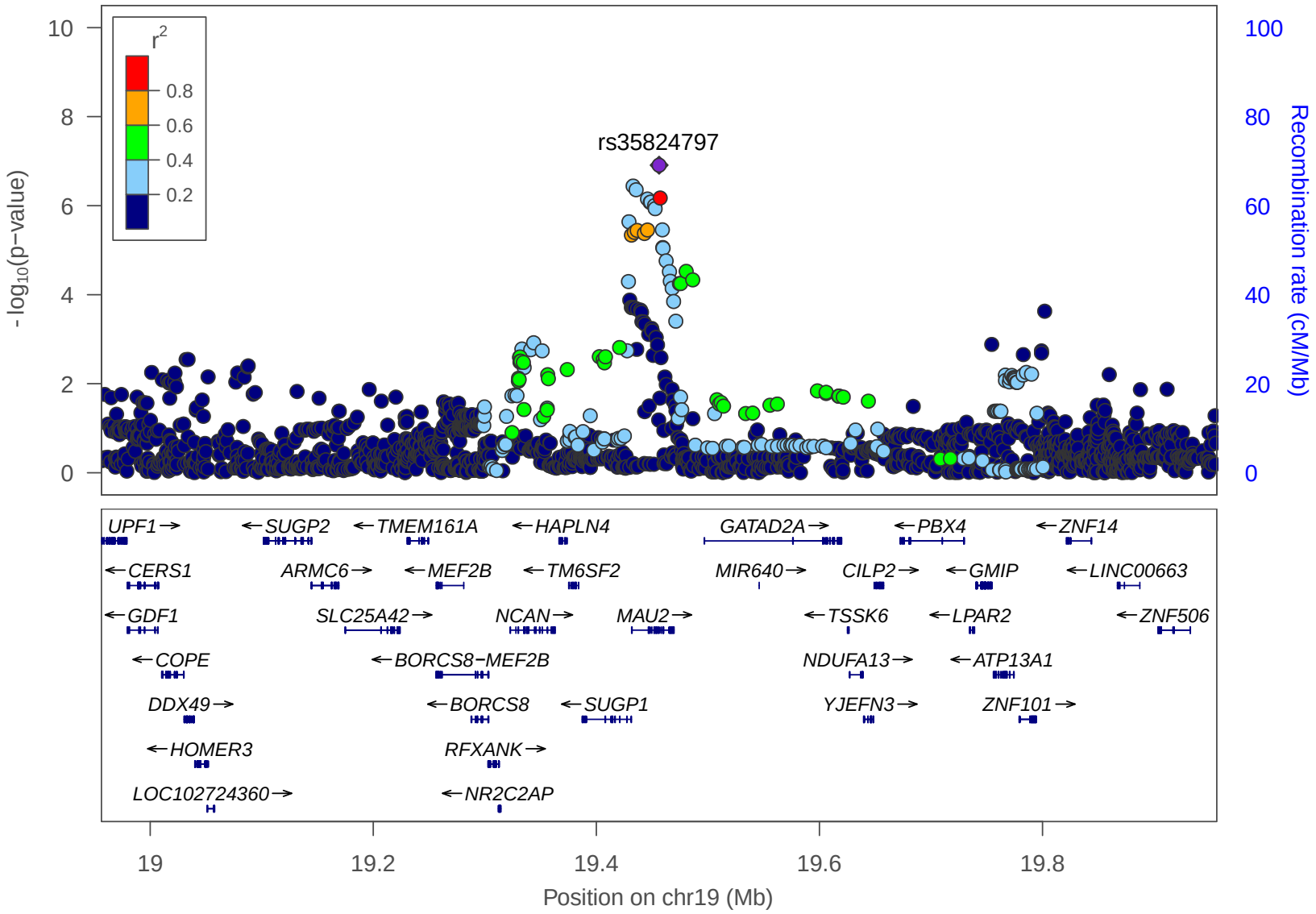
S5. Chromosome 16



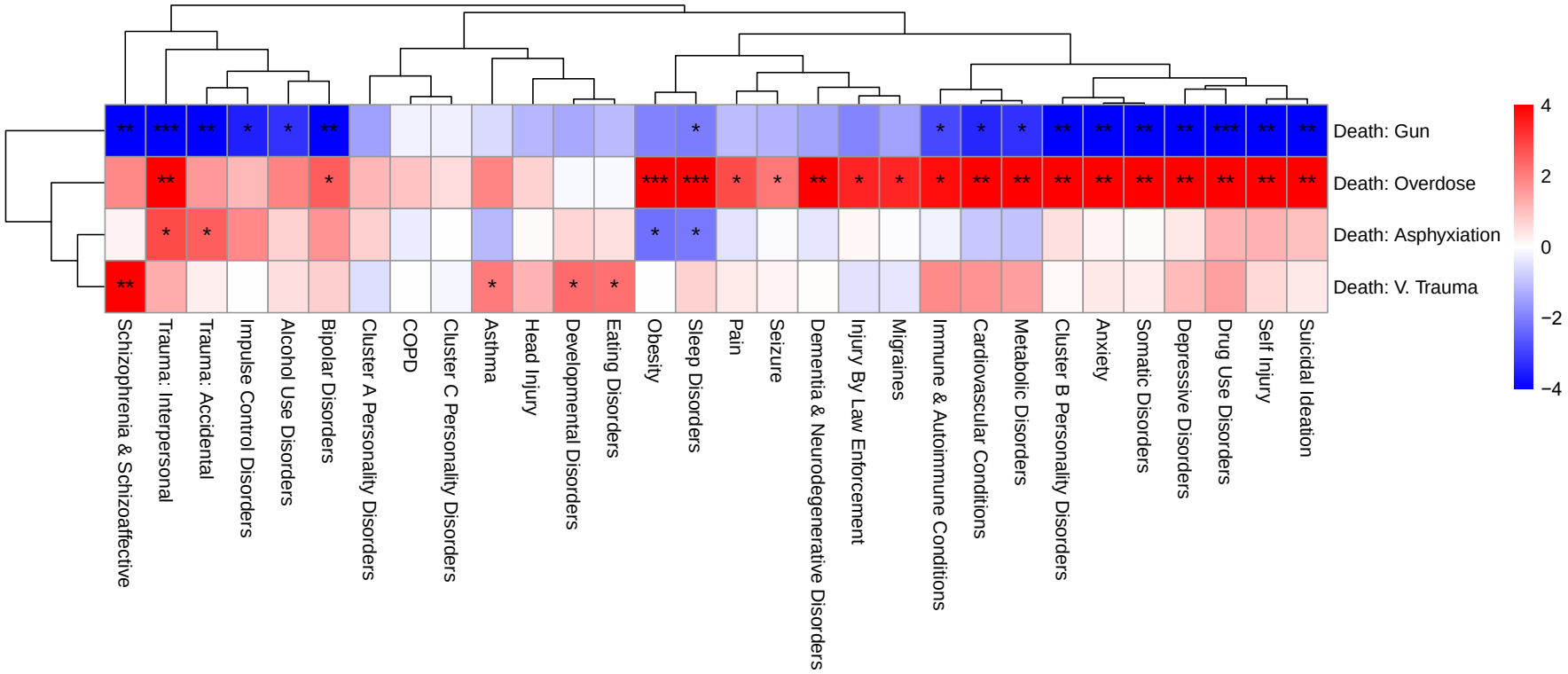
S6. Chromosome 17



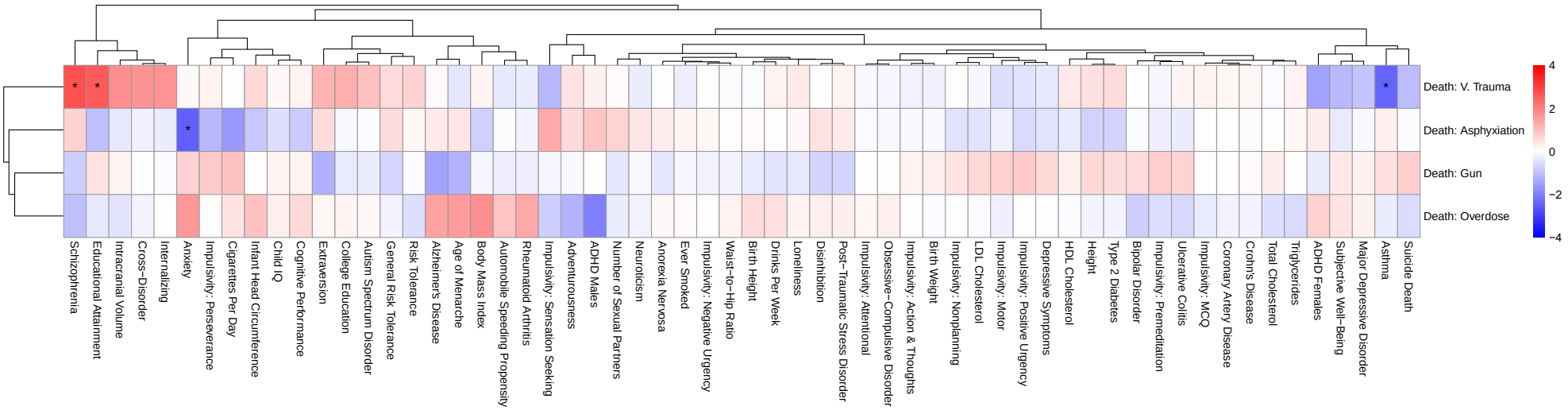
S7. Chromosome 17



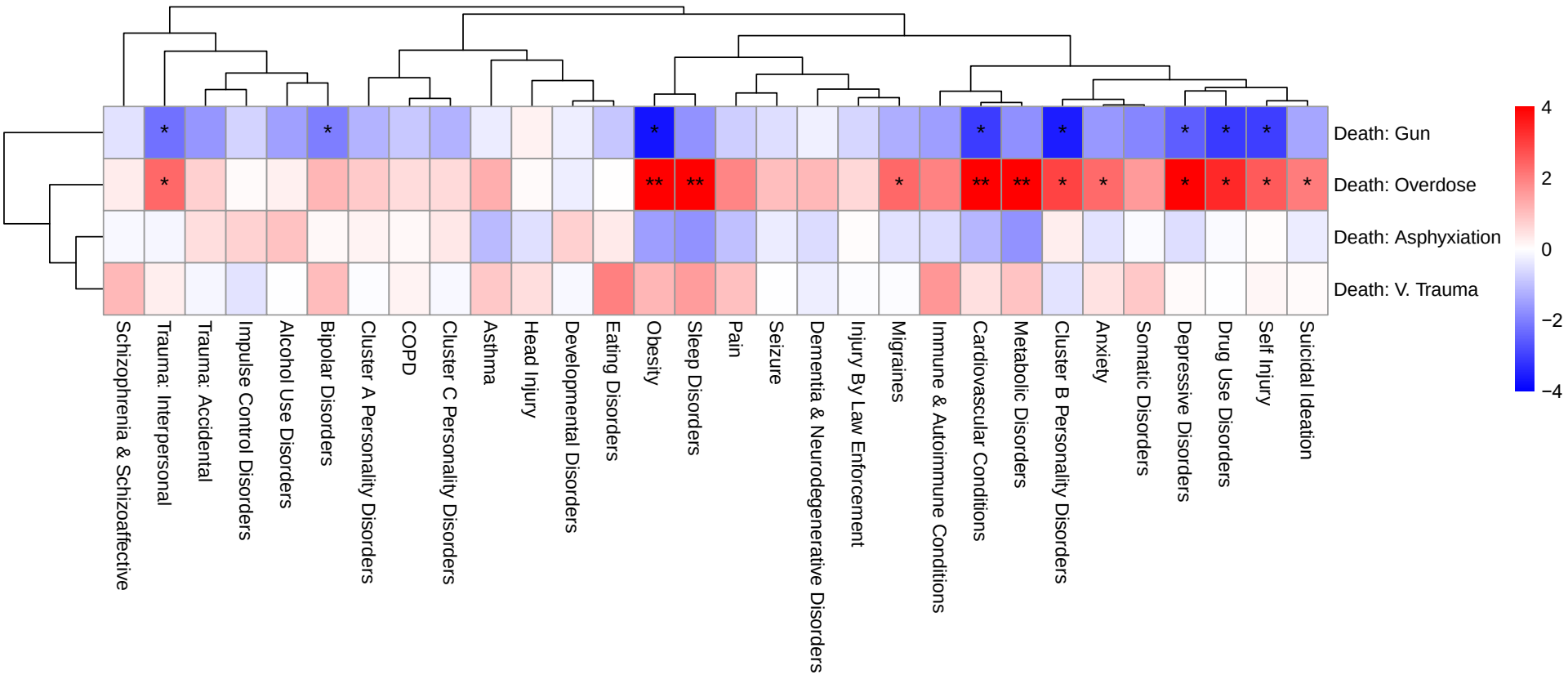
S8. Chromosome 19



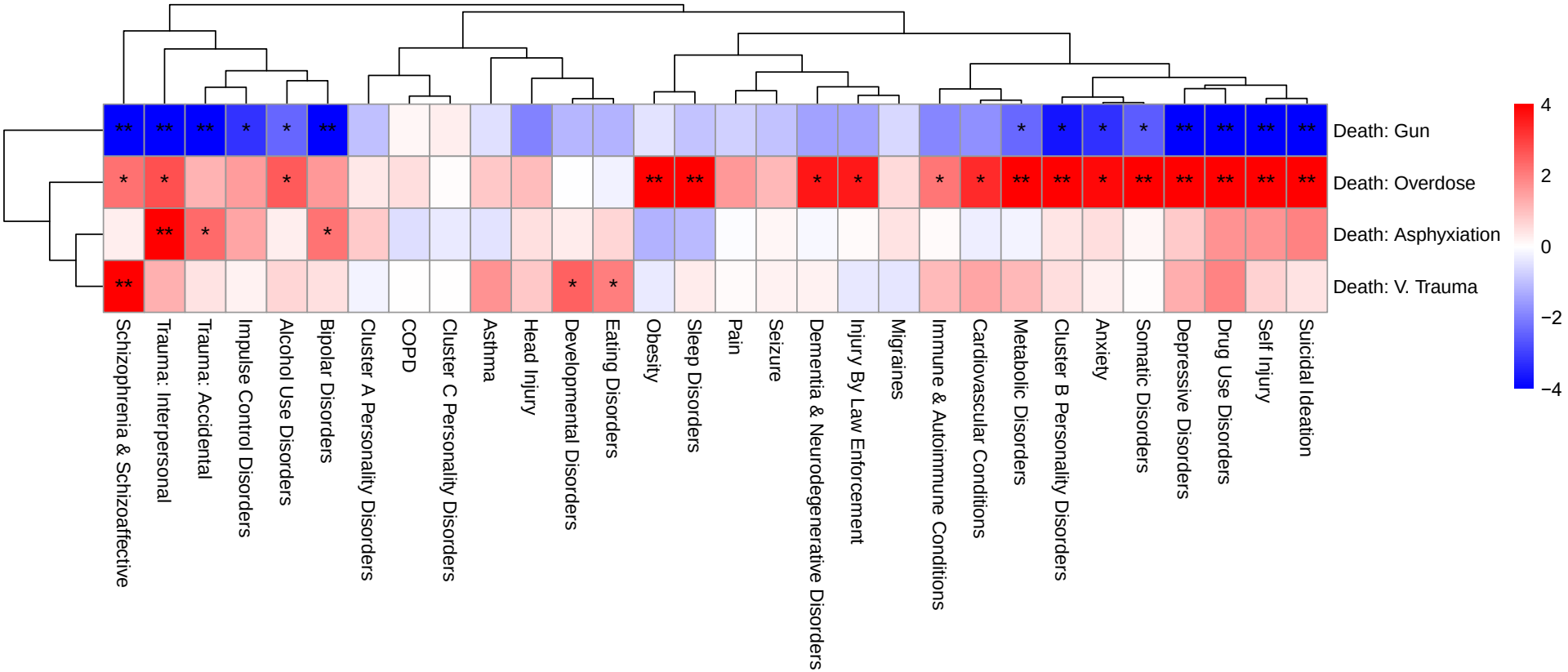
S9. Diagnostic Antecedents of Specific Mode of Suicide. Medical record diagnoses (x-axis) mapped to mode of death (y-axis) in the suicide death cases. Dendrograms based on k-means clustering of Euclidean distances are featured on both axes. Shading reflects the test statistic value as positive (red) or negative (blue). $p<10^{-2}$ (*), $p<10^{-4}$ (**), and $p<10^{-10}$ (***).



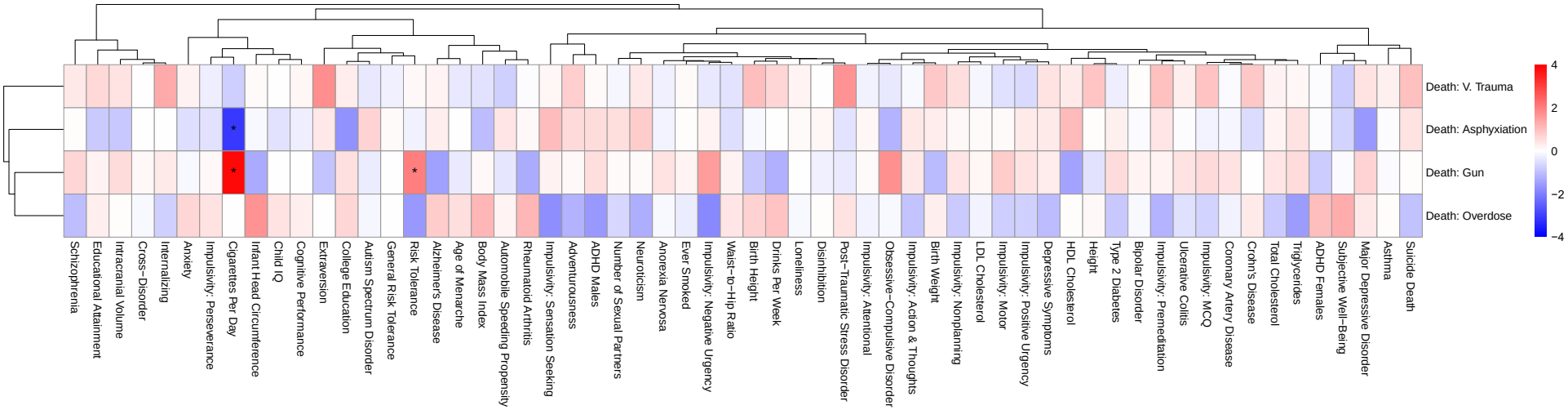
S10. PRS X MD: All Suicide Deaths



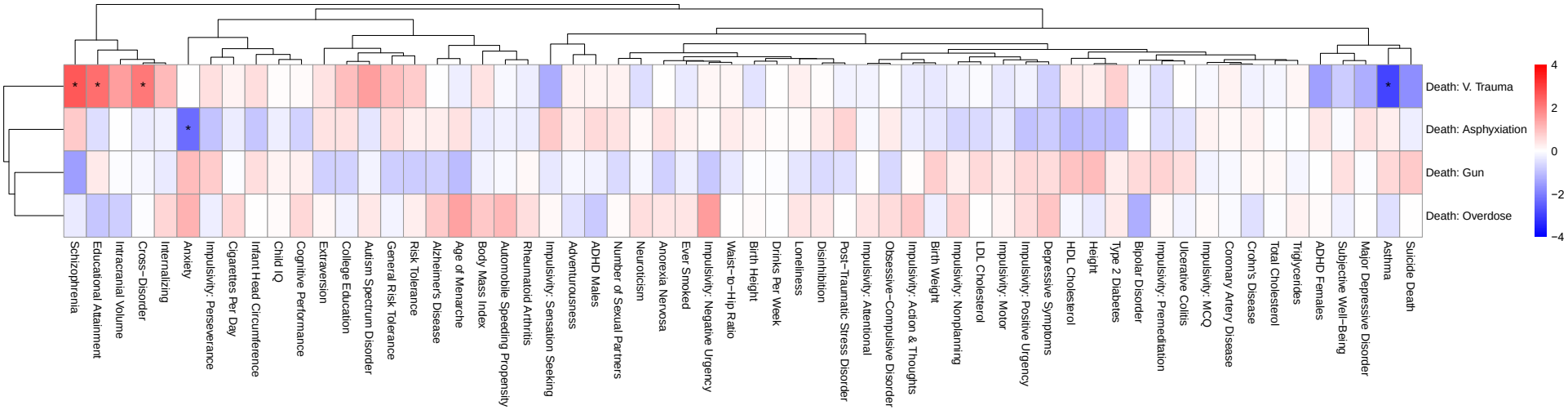
S11. ICD X MD Plots: Female Suicide Deaths > 25 Years Of Age



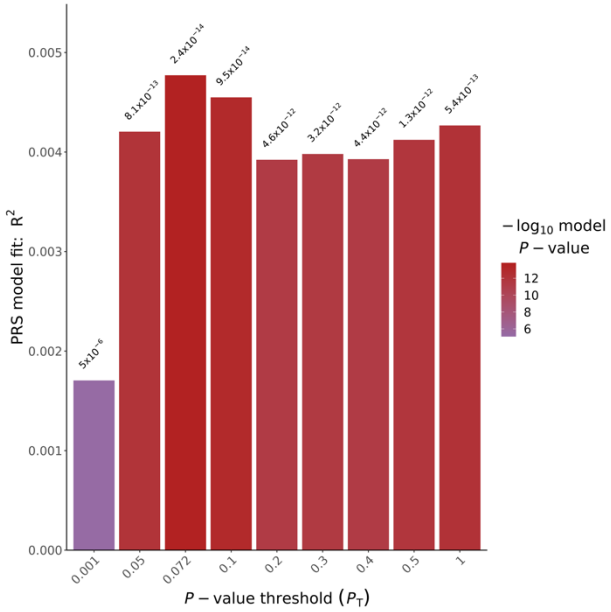
S12. ICD X MD Plots: Male Suicide Deaths > 25 Years Of Age



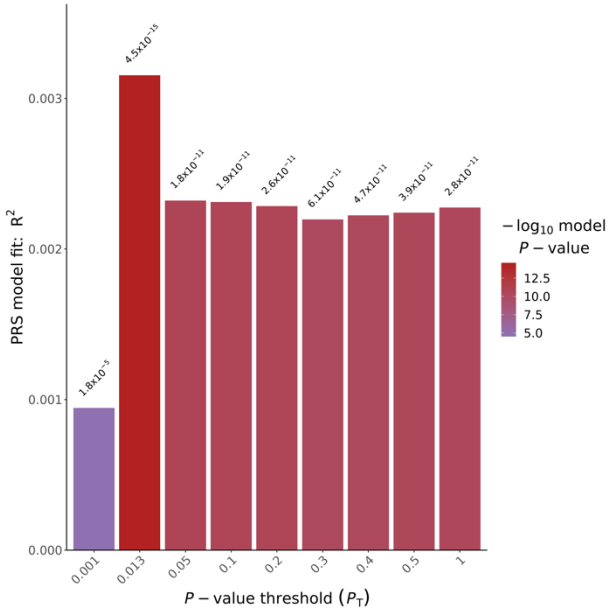
S13. PRS X MD: Female Suicide Deaths



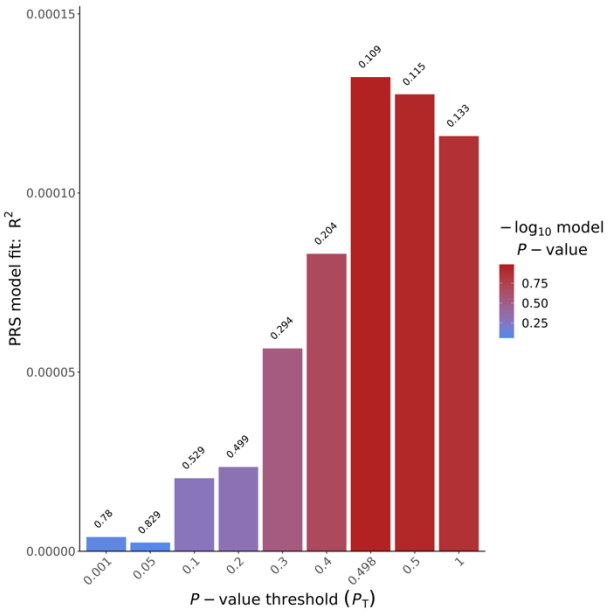
S14. PRS X MD: Male Suicide Deaths



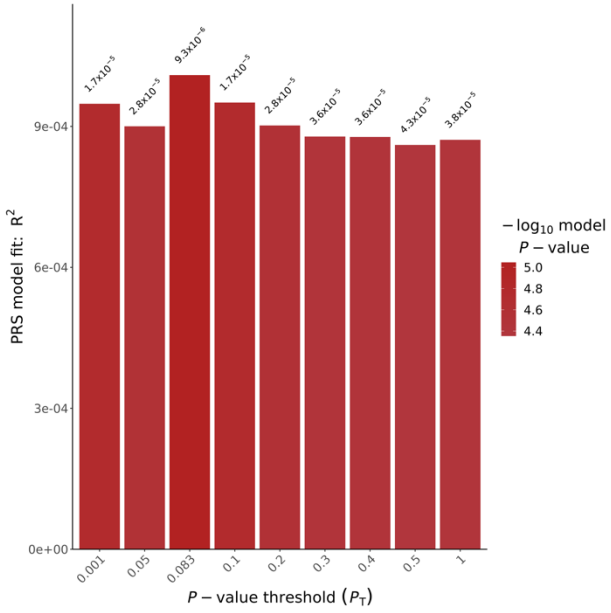
Autism Spectrum Disorder



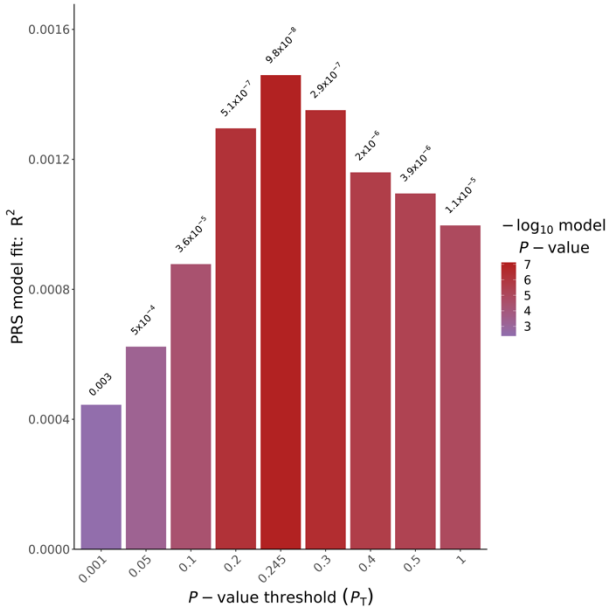
Depressive Symptoms



Loneliness



Alcohol: Drinks Per Week



Child IQ

S15. Plots of Additional Case-Control Differences in Polygenic Risks