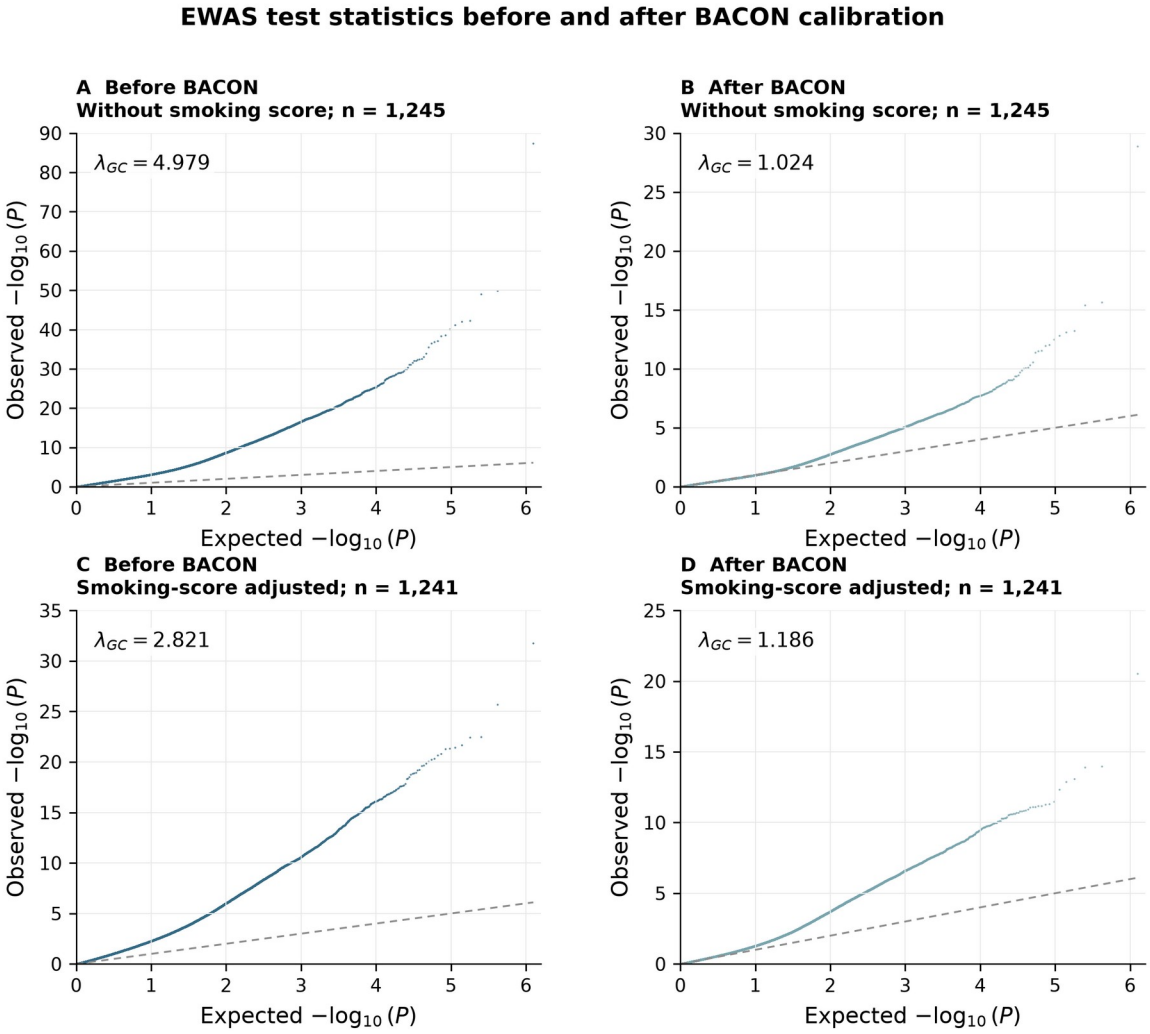
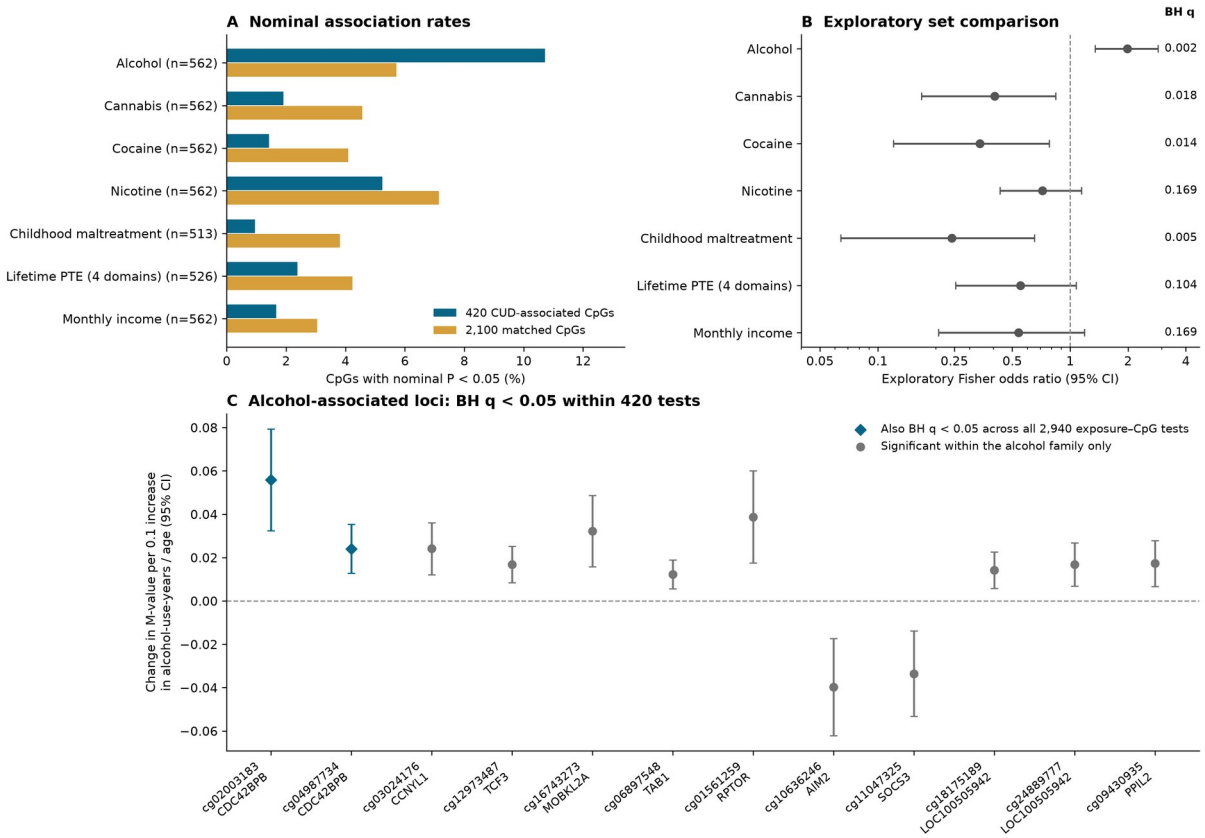
# Supplementary Figure 1



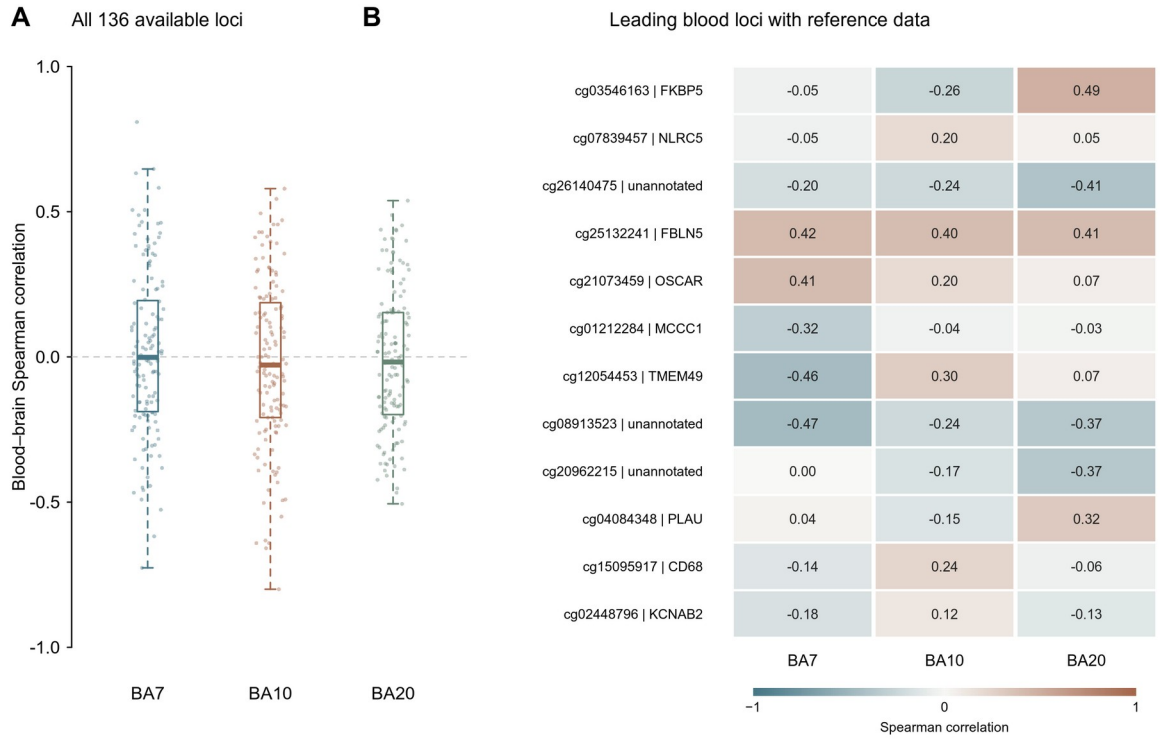

Q–Q plots before and after BACON calibration. Observed versus expected −log10(P) values are shown for all 627,130 CpGs in the model without the smoking score (A, B; n = 1,245) and the primary smoking-adjusted model (C, D; n = 1,241), before calibration (A, C) and afterward (B, D). Raw model P values were taken from the fitted EWAS; post-calibration P values were obtained from the BACON-calibrated statistics. Expected quantiles are based on (rank − 0.5)/627,130. The dashed diagonal denotes the null expectation. Genomic inflation factors (λGC), calculated as the median one-degree-of-freedom chi-square statistic implied by the P values divided by its null median, were 4.979 and 1.024 before and after BACON in the model without the smoking score, and 2.821 and 1.186 in the smoking-adjusted model. Panels use different vertical-axis limits to display their complete tails. BACON calibrates bias and inflation of the association statistics; it does not itself perform multiple-testing correction.

# Supplementary Figure 2




Within-CUD exposure associations at 420 smoking-adjusted EWAS loci and 2,100 matched comparison CpGs. (A) Proportions with nominal two-sided P < 0.05 from separate mixed models for each exposure; these are not counts of multiplicity-adjusted associations. (B) Exploratory between-set comparisons using Fisher conditional maximum-likelihood odds ratios and exact 95% confidence intervals, with BH correction across seven comparisons. These tests do not account for correlation among CpGs or matched sampling; their P values, intervals and BH values are descriptive exploratory summaries rather than a dependence-adjusted global enrichment test. (C) Twelve alcohol-associated loci meeting BH q < 0.05 within 420 tests. Effects are scaled per 0.1 increase in years of regular alcohol use divided by age. Diamonds identify the two loci also meeting BH q < 0.05 across all 2,940 tests spanning seven exposures. All models adjust for age, sex, five cell proportions, ten EPISTRUCTURE PCs, array position and methylation-derived smoking score, with a chip random intercept. Sample sizes are n = 562 for substance-use ratios and income, n = 513 for childhood maltreatment and n = 526 for lifetime PTE. Childhood maltreatment counts five severe CTQ domains, reconstructed from all 25 valid raw item responses with prescribed reverse scoring and thresholds of ≥13 for physical abuse, ≥13 for sexual abuse, ≥16 for emotional abuse, ≥13 for physical neglect and ≥18 for emotional neglect. Lifetime PTE counts four event categories with complete responses and is distinct from the combined seven-domain adversity index in the complementary PCA. No missing response was treated as absence of exposure. Comparators were matched 1:5 on mean and SD of beta-values without replacement and without using clinical association results. These analyses are exploratory within the discovery cohort. Complete CpG and complementary PCA results are reported in Supplementary Table 5.

## Supplementary Figure 3




Blood–brain reference methylation correlations for the final EWAS loci. All 420 smoking-adjusted blood EWAS CpG identifiers were queried in BECon on 19 September 2026; regional correlations were returned for 136, with the other 284 recorded as unavailable. (A) Spearman coefficients between blood and BA7, BA10 and BA20 methylation for all 136 available loci. Each point represents one CpG; horizontal box lines show the median, boxes the interquartile range and whiskers the most extreme observations within 1.5 interquartile ranges of the box. (B) The 12 highest-ranked blood EWAS loci with reference data, selected by blood EWAS P value rather than correlation magnitude. Colours share a scale from −1 to 1, and labels show coefficients rounded to two decimals. Gene labels use the first listed gene in the current EPIC annotation; all annotations are retained in Supplementary Table 9. BECon derives these correlations from paired blood and cortical samples from 16 individuals. They describe inter-individual covariation and are separate from the CUD–control analysis in independent BA9 donors. They do not establish cross-tissue replication. Full-precision coefficients and coverage of all 420 queried loci are provided in Supplementary Table 9.

## Supplementary tables

NATMH-25-1573A. Machine-readable tables are supplied as accompanying files. The complete discovery universe comprises 627,130 tested CpGs.

**Supplementary Table 1. Selected blood EWAS DMPs**

Supplementary_Table_1_DMPs.xlsx

**Supplementary Table 2. Complete summary statistics for both blood EWAS models**

Supplementary_Table_2_EWAS_smoking_adjusted.csv.gz and
Supplementary_Table_2_EWAS_without_smoking_score.csv.gz; see accompanying README

**Supplementary Table 3. Genomic-context enrichment and depletion**

Supplementary_Table_3_Genomic_context.xlsx

**Supplementary Table 4. Unified GO, KEGG and complementary Reactome gene-set analyses**

Supplementary_Table_4_GO_KEGG.xlsx; Supplementary_Table_4_Reactome.csv; see accompanying Reactome eligibility and provenance files

**Supplementary Table 5. Exposure associations, matched comparisons and exploratory PCA**

Supplementary_Table_5_Exposures.xlsx

**Supplementary Table 6. Alcohol-subgroup sensitivity analysis**

Supplementary_Table_6_Alcohol_subgroup.xlsx

**Supplementary Table 7. Postmortem BA9 analyses and cohort characteristics**

Supplementary_Table_7_Brain.xlsx

**Supplementary Table 8. Variance inflation factors**

Supplementary_Table_8_VIF.csv; see accompanying legend

**Supplementary Table 9. Reference comparisons and complete BECon query**

Supplementary_Table_9 folder; see ST9_Reference_comparisons_README.md and
ST9_BECon_README.md