Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Epigenome-wide association study in a large admixture cohort reveals immune and stress response epigenetic alterations in cocaine use disorder

Rodrigo Grassi-Oliveira

rogo@clin.au.dk

Pontifical Catholic University of Rio Grande do Sul https://orcid.org/0000-0001-9911-5921

Gabriel Fries

McGovern Medical School

Joy Schimidt

McGovern Medical School

Consuelo Walss-Bass

McGovern Medical School, Houston

Bruno Kluwe-Schiavon

McGovern Medical School, Houston

Thiago Viola

Pontifical Catholic University of Rio Grande do Sul

Diego Rovaris

University of São Paulo https://orcid.org/0000-0002-8910-4927

Eugenio Grevet

Universidade Federal do Rio Grande do Sul

Claiton Bau

Universidade Federal do Rio Grande do Sul

João Antônio Leal

Pontifical Catholic University of Rio Grande do Sul

Article

Keywords:

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9544880/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Title

Epigenome-wide association study in a large admixed cohort reveals immune and stress response epigenetic alterations in cocaine use disorder

Authors

Rodrigo Grassi-Oliveira[1,2]*, Gabriel R. Fries[3], Bruno Kluwe-Schiavon[1], Thiago Wendt Viola[1], Diego Luiz Rovaris[4], João Antônio Valadão Leal[1], Joao Eugenio Horacio Grevet[5], Claiton Henrique Dotto Bau[5], Joy Schmidt[1], and Consuelo Walss-Bass[3]

Affiliations

[1] Nanopsychiatry Lab, Brain Institute and School of Medicine, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, Brazil

[2] Translational Neuropsychiatry Unit, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[3] Department of Psychiatry and Behavioral Sciences, The University of Texas Health Science Center at Houston (UTHealth), Houston, TX, USA

[4] Department of Physiology and Biophysics, Institute of Biomedical Sciences, University of São Paulo (USP), São Paulo, Brazil

[5] Department of Psychiatry, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

Corresponding author

*Rodrigo Grassi-Oliveira, Ph.D.

Translational Neuropsychiatry Unit, Aarhus University, Aarhus, Denmark

and Brain Institute, PUCRS, Porto Alegre, Brazil

Email: rodrigo.grassi@pucrs.br; rogo@clin.au.dk

Manuscript type

Article

Word count (main text)

[Insert final word count once finalized]

Figures/Tables

[Insert final number once finalized]

Supplementary Information

Included

Conflict of interest statement

The authors declare no competing interests.

Data availability

All EWAS data supporting the findings of this study are available within the manuscript and Supplementary Information. Clinical data cannot be publicly shared due to ethical restrictions but are available from the corresponding author upon reasonable request and institutional approval.

Code availability

All analysis scripts used for EWAS preprocessing and statistical modeling are available from the corresponding author upon reasonable request.

Abstract

Cocaine use disorder (CUD) is frequently accompanied by tobacco use, polysubstance exposure and traumatic life events, complicating the interpretation of peripheral epigenetic signatures. We conducted a large epigenome-wide association study of CUD in peripheral blood DNA from 1,250 Brazilian adults, including 563 individuals with CUD and 687 non-using controls. After adjustment for technical factors, ancestry-related methylation structure, sex, age and estimated blood cell composition, 34 CpGs reached epigenome-wide significance ($p < 9 \times 10^{-8}$), and 79 top-ranked DMPs met FDR < 0.005. Most DMPs showed lower methylation in CUD and included loci annotated to *AHRR*, *FKBP5*, *SOCS3*, *RARA* and *ELMSAN1*. Enrichment analyses of hypomethylated DMP-associated genes converged on cytokine/interferon signaling, JAK–STAT regulation, platelet activation, endocytosis, focal adhesion and MAPK/Rap1/PI3K-Akt pathways, indicating a peripheral immune-inflammatory and cell-remodeling methylation profile. Within the CUD group, nicotine sensitivity analysis identified 78 nicotine-associated CpGs, only 12 of which overlapped with CUD-associated DMPs, suggesting that tobacco exposure explains a biologically expected subset, but not the majority, of the CUD methylation signal. Trauma burden and polysubstance-use histories explained partially distinct methylation components, with *FKBP5* linked to trauma-associated variance. Blood–brain analyses showed limited global portability but locus-specific convergence in postmortem BA9 for selected CpGs. These findings identify a predominantly hypomethylated peripheral methylation profile in CUD shaped by nicotine exposure, trauma, polysubstance use and immune-stress regulation.

Keywords

cocaine use disorder; epigenetics; DNA methylation; trauma; stress response; immune regulation

Introduction

Cocaine use disorder (CUD), characterized by the compulsive use of cocaine despite adverse medical, psychological, and behavioral consequences [1], remains a major global public health concern, with an estimated 23.5 million users worldwide. Global trends indicate a resurgence in stimulant drug use, including cocaine, with a 20% increase compared to a decade earlier [2]. Converging evidence highlights multilayered biological regulation in addiction, ranging from neuroanatomical alterations [3,4] to genetic vulnerabilities [5–7]. At the molecular level, epigenetic mechanisms—including DNA methylation, histone modifications, and non-coding RNAs—have been increasingly recognized as key contributors to the functional and structural brain changes associated with substance use disorders. Experimental models have shown that repeated cocaine exposure induces widespread DNA methylation changes in the brain and blood [8,9], reflecting both direct neurotoxic effects and systemic biological adaptations. In human studies, prior genome-wide methylation analyses have identified peripheral blood DNA methylation alterations associated with prenatal cocaine dependence [10,11] and in the umbilical cord blood from newborns exposed to cocaine prenatally [12], although these findings have been limited by modest sample sizes and minimal clinical characterization.

The role of potentially traumatic experiences (PTEs) has been consistently implicated in the initiation and maintenance of substance use disorders, supported by convergent evidence from large-scale epidemiological studies [13]. Childhood exposure to PTEs has been linked to an increased risk of cocaine use during adolescence [14], while cocaine dependence itself has been associated with higher rates of trauma exposure and comorbid post-traumatic stress disorder (PTSD) [15]. For instance, women with CUD exhibit higher rates of trauma-related disorders, greater social and occupational impairments, and more severe substance use than men [16]. In parallel, childhood trauma has been associated with an increased risk of suicide attempts among individuals with CUD, further underscoring the profound psychiatric burden linked to early adversity [17]. These associations are particularly salient in vulnerable populations where chronic violence and social inequities compound biological

susceptibilities [18]. In this context, the interplay between trauma exposure, cocaine use, and epigenetic dysregulation provides a compelling framework for understanding the multifactorial etiology of CUD.

Here, we sought to characterize DNA methylation alterations in peripheral blood associated with CUD in a large sample of admixture adults, and to evaluate their concordance with findings from postmortem prefrontal cortex tissue. We hypothesized that epigenetic dysregulation would be detectable in blood and associated with both CUD diagnosis and environmental exposures, providing mechanistic insights into the biology of addiction and informing the development of epigenetic biomarkers. Furthermore, by examining the impact of trauma exposure on the epigenetic landscape, we aimed to disentangle the interplay between environmental adversity and cocaine-related molecular alterations. This integrated approach is intended to advance our understanding of CUD pathophysiology and to guide the future development of targeted, personalized therapeutic strategies.

Results

Sample characteristics

Our study included 563 individuals with cocaine use disorder and 687 non-using controls. Cases were recruited from public detoxification centers dedicated to alcohol and drug rehabilitation in Southern Brazil. Inclusion criteria required participants to be between 18 and 45 years old, have a confirmed diagnosis of CUD according to the Structured Clinical Interview for DSM-5 (SCID-5). As expected, **Table 1** reveals marked sociodemographic and clinical differences between the control and CUD groups. Individuals with CUD had significantly lower income and educational attainment, as well as lower representation of White-identifying participants, compared to controls. Clinically, individuals with CUD exhibited high rates of lifetime substance use disorders, particularly nicotine (88.1%), cannabis (73.3%), and alcohol (68.7%) use disorders (**Supplemental Table 1**). In contrast, none of the control participants met criteria for a current SUD diagnosis; lifetime diagnoses were rare and reflected past substance use that no longer met diagnostic threshold at the time of assessment. Psychiatric comorbidities were also significantly more frequent in the CUD group. Depressive disorder was observed in 33.3% of individuals with CUD—compared to 23.0% in controls—and included 15.2% with substance-induced depression, absent among controls. Similarly, bipolar disorder was present in 7.4% of the CUD group versus 3.8% in controls, with 8.1% of CUD participants meeting criteria for substance-induced bipolar disorder, also not observed in the control group. These findings highlight the elevated burden of lifetime psychopathology and polysubstance involvement among individuals with cocaine use disorder, in contrast to the relative psychiatric stability of the control sample.

DNA methylation differences

We identified 34 differentially methylated CpG sites at $p < 9 \times 10^{-8}$ and 79 DMPs at FDR-adjusted $p < 5 \times 10^{-3}$ between individuals with CUD and controls (**Figure 1, Table 2**). Among the top 79 sites, 11 showed higher methylation and 68 showed lower methylation in individuals with CUD. In absolute annotation counts, DMPs were most frequently located in OpenSea regions, consistent with the broad representation of OpenSea probes on the EPIC array. However, enrichment analysis relative to the array background revealed a distinct pattern: DMPs were overrepresented in CpG island flanking regions, particularly S_Shore and S_Shelf, and depleted in CpG islands, whereas OpenSea regions showed slight relative depletion (**Figure 2**). DMPs were also depleted in promoter-proximal contexts, including TSS1500 and TSS200, with few sites overlapping 5′UTRs or exon boundaries (**Supplementary File 1**). The top-ranked sites are shown in Table 2, and the full list of DMPs is provided in **Supplementary File 1**, with the most significant annotated probes located in or near the aryl hydrocarbon receptor repressor (*AHRR*, cg05575921 and cg21161138), retinoic acid receptor alpha (*RARA*, cg17739917), suppressor of cytokine signaling 3 (*SOCS3*, cg13343932, cg11047325 and cg18181703), mitotic deacetylase associated SANT domain protein (*ELMSAN1*, cg02738868), the long non-coding RNA *AC068134.5* (cg01940273 and cg21566642) and FK506-binding protein 51 (*FKBP5*, cg03546163) genes.

Enrichment Analysis

Trait and disease enrichment of hypomethylated DMPs

Multi-database enrichment mapping of genes annotated to hypomethylated top-ranked DMPs in peripheral blood from individuals with CUD converged on an immune-inflammatory and signal-transduction profile. Across Enrichr-based pathway resources, including Reactome, BioPlanet, and MSigDB Hallmarks, enriched terms highlighted cytokine-driven pathways, including IL-2, IL-6, canonical JAK–STAT signaling, interferon responses, and broader immune-system activation. Reactome terms further implicated interleukin signaling, interferon-α/β signaling, platelet activation and hemostasis, G-protein βγ signaling through PI3Kγ, thrombin signaling through proteinase-activated receptors, and growth hormone receptor signaling. Additional enrichment for SUMOylation and retinoic-acid-associated transcriptional regulation suggested that hypomethylated DMPs in CUD also map to regulatory programs involved in post-translational modification and cellular differentiation (**Supplementary File 2.1**).

Disease, trait, transcription-factor, and perturbation libraries provided convergent support for this interpretation. HDSigDB and GWAS-based resources implicated myeloid and microglia-related signatures, eosinophil traits, and reduced lung-function loci, whereas disease resources including DisGeNET, Orphanet, and ClinVar highlighted leukocyte-adhesion defects, hyper-IgE syndrome, acute myeloid leukemia, pneumonia, tuberculosis susceptibility, and Alzheimer's disease. Transcription-factor enrichment further supported an immune-regulatory architecture, with enrichment for hematopoietic and inflammatory regulators including KDM2B, MYB, BRD4, STAT4/STAT6, and PU.1. VirusMINT interactome analysis identified selected hypomethylated-DMP-associated genes as viral-interaction targets, including NPM1, HNRNPA1, and FKBP5. GO analyses revealed significant enrichment for processes related to immune cell differentiation, including myeloid and lymphocyte differentiation, cell-cycle regulation, small GTPase-mediated signal transduction, phospholipid and glycerophospholipid biosynthesis, and intrinsic apoptotic signaling. Collectively, these multi-layered enrichments suggest that CUD-related hypomethylation preferentially maps to immune, cytokine-responsive, hemostatic, and antiviral effector loci, consistent with a peripheral inflammatory and host-defense epigenetic signature (**Supplementary File 2.1**). In line with these observations, KEGG pathway enrichment indicated involvement of MAPK, Rap1, PI3K-Akt, and AMPK signaling, alongside focal adhesion, platelet activation, and endocytosis pathways (**Figure 3**).

Protein–protein interaction network of hypomethylated DMP-associated genes

To assess whether genes annotated to hypomethylated DMPs formed a functionally coherent module, we submitted the 45 annotated genes to STRING v12 using a medium interaction confidence threshold of 0.4. The resulting network comprised 45 nodes connected by 29 edges, exceeding the 13 edges expected by chance and indicating significant PPI enrichment ($p = 7.98 \times 10^{-5}$; **Figure 4**). This excess connectivity suggests that CUD-associated hypomethylation maps to an inter-related protein network rather than being randomly distributed across functionally unrelated genes. Two functionally interpretable subnetworks were apparent. The first involved innate immune and JAK–STAT-related genes, including *SOCS3, CISH, BST2, TRIM69, PTPRJ, CSF2RB, TNFSF10,* and *PLAU*, consistent with the interferon and cytokine-signaling signatures recovered in pathway analyses. The second involved cytoskeleton, adhesion, and remodeling-related genes, including *SEMA7A, ASAP1, SH3PXD2A, FERMT3, ECE1, FUT7,* and *PTPRJ*, implicating pathways relevant to leukocyte adhesion and cell trafficking. Peripheral regulatory genes, including *HDAC4, HMGA2, FKBP5,* and *AHRR*, may represent exposure-sensitive regulatory nodes linking methylation variation with immune transcriptional programs.

Trait and disease enrichment of hypermethylated DMPs

Trait-enrichment mapping of genes annotated to the small set of probes with increased methylation in peripheral blood DNA from individuals with CUD suggested two partially convergent annotation axes. First, a connective-tissue and vascular axis was largely driven by FBLN5, which was significantly enriched in ClinVar terms related to autosomal recessive cutis laxa and age-related macular degeneration, with additional nominal

DisGeNET annotations involving arterial, valvular, and extracellular-matrix-related phenotypes. Second, a neurodevelopmental and neuromuscular axis emerged across DisGeNET and Orphanet annotations, involving KCNAB2, MYH10, MYO1G, and MKL2, with terms related to late fontanel closure, intellectual disability, epileptic phenotypes, cerebellar ataxias, Charcot–Marie–Tooth disease, and childhood-onset proximal spinal muscular atrophy. Together, these findings indicate that hypermethylated DMPs in CUD map to a narrow gene set enriched for disease annotations related to extracellular-matrix integrity, vascular and connective-tissue phenotypes, and neurodevelopmental or neuromuscular disorders. Given the small number of hypermethylated probes and the annotation-driven nature of the analysis, these results should be interpreted as hypothesis-generating rather than evidence of direct functional impairment (**Supplementary File 2.2**). No significant PPI enrichment was detected for genes annotated to hypermethylated DMPs.

Effects of trauma exposure and polysubstance use on differentially methylated CpG sites

Group-level comparisons revealed differences between individuals with CUD and controls in both lifetime trauma exposure (**Table 3**) and patterns of substance use (**Table 1**), particularly the extent and diversity of regular drug use. Among CUD participants, reports of life-threatening experiences, sexual and physical violence, and severe childhood adversity were markedly elevated. Additionally, chronic exposure to multiple substances—cannabis, alcohol, nicotine, and cocaine—was frequent, supporting a polysubstance profile. Notably, 89% of participants in the CUD group reported at least one PTE, compared to only 20% of controls. To determine whether these two dimensions—trauma burden and polysubstance exposure—jointly shape the methylation landscape in CUD, we conducted a principal component analysis (PCA) on the top 79 differentially methylated CpGs identified in the full sample (**Supplementary File 3**). This dimensionality reduction yielded 11 components with eigenvalues > 1, cumulatively explaining 76.3% of total variance (KMO = 0.973; **Figure 5A**). Component scores were extracted for each participant with CUD and used as outcomes in stepwise regression models. Predictors included the number of PTEs, proportion of lifetime years of regular cannabis, alcohol, nicotine, and cocaine use, as well as monthly income (used as a proxy for socioeconomic status).

The resulting models demonstrated that both trauma exposure and polysubstance use contributed independently and interactively to epigenetic variance across multiple components. For example, PC1 was predicted by regular cannabis use; PC2 by nicotine and cannabis use alongside trauma; and PC4, PC5, PC10, and PC11 by regular cocaine use. Alcohol use predicted PC6. Trauma exposure independently predicted PC2, PC4, PC7, PC8, and PC10. Notably, the *FKBP5*-associated CpG cg03546163 loaded exclusively on PC10, the component most strongly linked to PTE burden. Across all PCs, cannabis, alcohol, and trauma were the most frequent and consistent predictors. For PC 1, only years of regular cannabis use emerged as a significant predictor (**Figure 5**). For PC 2, both years of regular nicotine/cannabis use, and PTE count were significant. Years of regular cocaine use predicted PC 4 as well as PCs 5, 10, and 11, whereas PTE count also predicted PCs 4, 7, 8, and 10. Interestingly, the *FKBP5*-associated CpG site cg03546163 loaded exclusively on PC 10 (**Supplementary File 3**), the component for which PTE count showed its strongest predictive effect in the regression analysis. Overall, years of regular cannabis and alcohol use, as well as PTE, were the most frequent predictors, reaching significance in at least five of the eleven PCs. Monthly income, included as a proxy for socioeconomic status within the CUD group, did not emerge as a significant predictor in any of the regression models. Thus, diverse substance-use histories extending beyond cocaine, and indicative of a polysubstance pattern effect, explained a substantial portion of the epigenetic variation within the CUD group. However, trauma burden accounted for additional, non-overlapping methylation variance that substance use could not fully capture.

To further characterize the biological relevance of these methylation patterns, we performed gene set enrichment analyses on each of the top components (PC1–PC4) principal component (**Supplementary File 4)**. For each PC, CpG sites with absolute loadings ≥ 0.4 were mapped to genes using EPIC array annotations, and enrichment was conducted using Gene Ontology (GO), KEGG, DisGeNET, and curated molecular signatures (MSigDB C2). PC1 showed minimal biological coherence, with only one significant gene set (related to mitotic arrest by docetaxel) and no disease enrichment. PC2 displayed robust enrichment for neuroendocrine and immune-metabolic pathways, including JAK-STAT, prolactin, leptin, insulin, and interferon signaling,

implicating a convergence of stress and metabolic regulation. PC3 exhibited strong associations with hematological malignancies, inflammatory processes, and GPCR-mediated signaling, pointing toward dysregulated immune and vascular responses. PC4 was the only component with significant pathway-level enrichment in both KEGG and GO databases, implicating structural remodeling processes including tight junction integrity, actin cytoskeleton regulation, motor protein transport, and prolactin signaling. Disease enrichment analyses highlighted diverse developmental and inflammatory phenotypes, including congenital anomalies, intrauterine infections, respiratory distress, and vascular malformations.

Associations between DMPs and clinical variables

Results are shown in **Supplementary File 5**. Within peripheral blood DNA of individuals with CUD, a hypomethylated antiviral/cytokine module (including *SOCS3, TRIM69, FKBP5, ELMSAN1*) displayed small but coherent positive correlations with lifetime anxiety-spectrum disorders, whereas the same CpGs showed negative or null associations with depressive phenotypes. This pattern reinforces the interpretation that cocaine leaves an immune–stress epigenetic scar more closely aligned with anxiety liability than with depression. As expected, smoking-related loci such as *AHRR* (cg05575921) and F2RL3 (cg03636183) emerged among the top EWAS hits ($\rho \approx -0.28$ to -0.38), consistent with their role as canonical biomarkers of tobacco exposure and their tendency to dominate EWAS signals in substance-use cohorts [19,20]. Notably, *SOCS3* (cg11047325, cg13343932, cg18181703) exhibited replicated, directionally consistent correlations with multiple clinical domains (family, legal, psychiatric) as well as with the number of SUD diagnoses. In parallel, *FKBP5* (a key regulator of stress response and glucocorticoid receptor signaling) and *AHRR* both showed negative trends with composite medical and psychiatric scores, consistent with the involvement of chronic stress and inflammation pathways in addiction [21,22]. Importantly, *SOCS3, FKBP5*, and an intergenic CpG (cg01940273) within the ALPPL2–CHRND locus on chromosome 2 remained robustly correlated with polysubstance severity even after nicotine phenotypes were excluded, suggesting that these associations are not solely driven by smoking.

Exploring the Link Between Peripheral Blood and Brain Epigenetics

Given the known and often observed tissue specificity of many methylation sites, we initially examined the top blood-identified CpGs in the BECon web application [23] to explore potential blood–brain correlations. This in silico analysis indicated that only a minority of the significant CpGs displayed moderate positive correlations (up to $r \approx 0.45$–0.50) between blood and prefrontal cortex (BA7, BA10, BA20), with most loci showing weak or even inverse correlations. As expected, the canonical smoking markers *AHRR* and *F2RL3* exhibited strong blood-cell composition effects and low cross-tissue portability, whereas stress- and immune-related loci such as *FKBP5* and *SOCS3* demonstrated more consistent positive correlations with cortical methylation (**Supplementary File 6)**.

Our analysis of postmortem brain tissue identified 18 DMPs (11 hypomethylated and 7 hypermethylated) under a lenient threshold of FDR < 0.2 (**Supplementary File 7)**. Convergent hypomethylation of SEMA7A and *ASAP1* (cell migration), *PTPRJ* (immune signaling), and *SHMT2* (mitochondrial one-carbon metabolism) in both cocaine-dependent brain and peripheral blood underscores the translational coherence of our EWAS results, whereas the persistent signal at the smoking sentinel *F2RL3* highlights residual tobacco-related confounding. Several brain-specific hypermethylated loci—notably *TMEM49* (cg12054453), *RPS6KA5*, and a doublet at *PIK3R5*—also displayed nominal shifts in blood, while others such as *LOC100996291* and cg00841849/*FBLN5* appeared restricted to brain tissue. This pattern suggests that autophagy (*TMEM49*), PI3K–MAPK signaling (*RPS6KA5/PIK3R5*), and related stress-response pathways are epigenetically disrupted both centrally and peripherally in CUD, whereas extracellular-matrix alterations may represent brain-specific changes. When applying a stringent threshold of FDR < 0.05, only four loci—*ASAP1*, *MKL2, SEMA7A*, and *TMEM49*—remained significant, underscoring their robustness as brain-associated methylation signals in CUD.

Sensitivity analysis

Because nicotine use was highly prevalent among individuals with CUD [24,25] and hypomethylation of the *AHRR* locus—the top-ranked finding in our EWAS—has been consistently reported as a robust biomarker of smoking [26–28], we conducted a sensitivity EWAS restricted to the CUD group using lifetime nicotine exposure adjusted for age as the main predictor. This analysis identified 78 nicotine-associated CpGs at FDR < 0.005 (Supplementary File 8). Twelve of these CpGs overlapped with the CUD-associated DMPs from the primary EWAS, and 11 showed concordant directional effects. The overlapping loci included canonical smoking-sensitive CpGs annotated to *AHRR* and *F2RL3*, as well as CpGs annotated to *PRSS23*, *RARA*, *ELMSAN1*, *GNG12*, and *MYO1G*. Thus, nicotine exposure explained a limited but biologically expected subset of the CUD-associated methylation signal, whereas most CUD-related DMPs were not recapitulated by lifetime nicotine exposure within the CUD group.

Discussion

This large EWAS of cocaine use disorder identifies a predominantly hypomethylated peripheral methylation signature that converges on immune-inflammatory, cytokine/interferon, stress-regulatory and cell-remodeling pathways. The study advances the field in three ways. First, it shows that the strongest peripheral methylation signal in CUD includes canonical smoking-sensitive loci, but extends beyond them to genes involved in glucocorticoid signaling, cytokine regulation, retinoid biology, chromatin regulation, adhesion and cellular trafficking. Second, by modelling nicotine exposure, polysubstance-use histories and potentially traumatic events within the CUD group, the analysis shows that peripheral methylation variation in CUD reflects a layered environmental architecture rather than a single exposure signal. Third, the integration with BECon and postmortem BA9 data indicates limited global blood–brain portability, but suggests locus-specific peripheral–central convergence for a subset of CpGs. Together, these findings position CUD as a systemic disorder with a measurable peripheral epigenetic footprint shaped by tobacco exposure, trauma burden, polysubstance use and immune-stress regulation.

Previous work has implicated epigenetic dysregulation in the transition from cocaine exposure to compulsive use and neurobiological toxicity [29], but available EWAS studies have generally been limited by modest sample sizes and by the difficulty of separating cocaine-associated signals from the heavy burden of tobacco exposure, polysubstance use, and trauma that commonly accompanies CUD [10,11]. By leveraging a substantially larger and deeply phenotyped cohort, our study provides greater statistical power and clinical resolution to examine these overlapping sources of methylation variation. We identified 79 top-ranked DMPs distributed across multiple chromosomes, with a predominantly hypomethylated profile that mapped to coherent immune-cytokine, interferon, stress-regulatory, cytoskeletal, adhesion, and cell-trafficking pathways. Beyond defining this peripheral methylation landscape, we related DMP-derived components to clinically relevant dimensions within the CUD group, including trauma burden and lifetime polysubstance exposure. This integrative approach highlights a set of biologically plausible candidate loci, including *FKBP5, SOCS3, RARA, ELMSAN1, SEMA7A,* and *ASAP1*, while emphasizing that CUD-associated methylation reflects a layered exposomic signature rather than a cocaine-only effect.

The strongest differentially methylated CpG in our EWAS was cg05575921, located in intron 1 of *AHRR*, which showed markedly lower methylation in individuals with CUD. This locus is one of the most consistently replicated smoking-sensitive CpGs [27,28,30,31], and its prominence in our analysis reinforces the need to explicitly account for tobacco exposure in substance-use EWAS. Prior studies have linked reduced methylation at cg05575921 to altered *AHRR* expression in peripheral blood mononuclear cells [32], but transcriptomic data would be required to determine whether this relationship is present in our cohort. Functionally, *AHRR* modulates aryl hydrocarbon receptor signaling (AhR), a pathway involved in xenobiotic sensing, immune regulation and inflammatory responses [33,34]. Thus, reduced methylation at *AHRR* in CUD likely reflects, at least in part, the high burden of tobacco exposure in this population, while also marking a broader immune–xenobiotic response axis relevant to substance-use phenotypes. Consistent with this broader interpretation, diminished cg05575921

methylation has also been reported in other stress- and substance-related cohorts, including cannabis use disorder and PTSD [32].

CUD in our cohort was associated with a clear immune-inflammatory methylation profile. Enrichment analyses of genes annotated to hypomethylated DMPs highlighted cytokine and interferon signaling, JAK–STAT-related regulation, platelet activation and hemostasis, endocytosis, focal adhesion, and MAPK/Rap1/PI3K-Akt/AMPK signaling, consistent with clinical and experimental evidence implicating systemic inflammation in cocaine use [35–37]. Key contributors to this profile included DMPs annotated to *SOCS3* [38], *NLRC5* [39], *RARA* [40], and *FKBP5* [41]. The presence of *SOCS3* among the top-ranked DMPs is particularly relevant given its role as a negative regulator of cytokine and JAK–STAT signaling, while *RARA* and *NLRC5* implicate retinoid-linked immune regulation and antigen-presentation pathways. In parallel, reduced methylation at *FKBP5* cg03546163 is consistent with prior EWAS evidence linking this locus to substance exposure [11]. Previous work has associated lower methylation at this site with altered *FKBP5* transcription [42] and glucocorticoid receptor regulation [43], although transcriptomic and endocrine data would be required to determine whether this mechanism operates in the present cohort. Reduced methylation at *FKBP5* cg03546163 has also been reported in relation to early-life trauma [44], providing external support for the link between this locus and stress-related biological embedding. In our cohort, lifetime exposure to potentially traumatic events was markedly higher in individuals with CUD than in controls, and PCA of the 79 CUD-associated CpGs yielded 11 components explaining 76.3% of the methylation variance. In stepwise regression models restricted to CUD cases, polysubstance-use metrics, particularly lifetime cannabis and alcohol exposure, were frequent predictors of methylation-derived components. However, PTE count explained additional, non-overlapping variance across multiple PCs. The *FKBP5*-associated CpG cg03546163 loaded exclusively on PC10, the component for which PTE count showed its strongest predictive effect. These findings support a model in which immune-stress methylation signatures in CUD are shaped not only by drug exposure, but also by cumulative traumatic experiences.

The convergence of hypomethylated DMPs on immune, cytokine-responsive and host-defense annotations suggests that CUD is accompanied by a peripheral epigenetic state resembling chronic immune priming rather than a narrow drug-specific molecular lesion. Cocaine exposure occurs in a clinical context marked by tobacco use, recurrent infections, pulmonary stress, sleep disruption, psychosocial adversity and repeated activation of stress systems, all of which may converge on inflammatory and myeloid regulatory programs [45–47]. The enrichment of leukocyte-adhesion, interferon, cytokine and viral-response annotations should therefore be interpreted less as evidence for specific infectious or hematological diagnoses and more as evidence that the methylation signal is concentrated in genes that regulate immune surveillance, cellular trafficking and host-defense readiness. In this sense, the EWAS supports a model in which CUD is biologically embedded in peripheral immune regulation, consistent with clinical observations that cocaine dependence is associated with systemic inflammatory burden and medical vulnerability. This interpretation is strengthened by independent immune phenotyping work showing that CUD is associated with altered adaptive immune-cell maturation, particularly in individuals with greater withdrawal severity [47]. Although those cellular findings and the present methylation results derive from different analytical layers, they converge on the same biological theme: cocaine dependence is accompanied by immune remodeling that is not captured by psychiatric diagnosis alone. The methylation data extend this observation by suggesting that immune dysregulation in CUD may be encoded not only in circulating cell proportions, but also in regulatory marks at cytokine, interferon, adhesion and stress-responsive loci.

Given the tissue-specific nature of DNA methylation, the blood–brain analyses provide an important boundary condition for interpreting the peripheral EWAS findings. BECon analyses indicated limited global portability of most blood-identified CpGs, particularly for canonical smoking-sensitive loci such as *AHRR* and *F2RL3*. Nevertheless, selected loci showed evidence of cross-tissue convergence. In postmortem BA9, four blood-identified DMPs—*SEMA7A* cg00310412, *ASAP1* cg10068788, *PTPRJ* cg25917651 and *SHMT2* cg16323416—showed directionally concordant hypomethylation under a lenient FDR < 0.20 threshold. These genes are involved in processes relevant to cell migration, axon guidance [48–50], immune signaling and

mitochondrial one-carbon metabolism, suggesting that a subset of peripheral methylation changes may overlap with brain-associated regulatory alterations in CUD. Under a stricter FDR < 0.05 threshold, *ASAP1, MKL2, SEMA7A* and *TMEM49* remained significant in BA9, further supporting locus-specific—but not global—blood–brain convergence. Conversely, several loci showed tissue-restricted or directionally distinct patterns, including PI3K–MAPK/autophagy-related genes such as *TMEM49, RPS6KA5 and PIK3R5*, and the persistence of F2RL3 reinforces that residual tobacco-related effects cannot be fully excluded.

To our knowledge, this is the largest epigenome-wide study to date in CUD and the first to integrate (i) polysubstance-adjusted blood EWAS, (ii) trauma-derived PCs, and (iii) targeted replication in prefrontal cortex. However several limitations should be considered. First, the cross-sectional design precludes causal inference. The methylation differences identified here may reflect chronic cocaine exposure, polysubstance use, tobacco exposure, trauma-related biological embedding, pre-existing vulnerability [51], socioeconomic adversity, or interactions among these factors. Therefore, the present findings should not be interpreted as evidence that cocaine exposure alone caused the observed methylation changes. Longitudinal cohorts, genetically informed designs, and experimental cellular models will be required to distinguish antecedent vulnerability from exposure-related methylation remodeling. Second, polysubstance use was highly prevalent in the CUD group. Although we explicitly examined nicotine, cannabis, alcohol and cocaine exposure, and performed a within-CUD nicotine sensitivity analysis, residual confounding by tobacco and other substances remains likely, particularly for canonical smoking-sensitive loci such as *AHRR* and *F2RL3*. Third, key lifestyle and clinical covariates with known effects on DNA methylation, including sleep cycle, diet, physical activity were unavailable or incompletely measured. Fourth, methylation was measured in whole blood, and although estimated blood cell proportions were included as covariates, residual cell-composition effects cannot be excluded. Fifth, the postmortem brain analysis was modest in size and restricted to BA9, limiting statistical power and regional generalizability. Accordingly, the blood–brain findings should be interpreted as preliminary evidence of locus-specific convergence rather than broad validation of the peripheral EWAS signature. Sixth, enrichment analyses were based on genes annotated to top-ranked DMPs and are therefore exploratory and subject to annotation bias, including unequal CpG coverage across genes and variable database curation. Finally, although the Brazilian cohort increases ancestral and socioeconomic diversity in CUD epigenetics, replication in independent populations will be essential to establish generalizability.

In summary, this large, deeply phenotyped EWAS identifies a predominantly hypomethylated peripheral methylation profile associated with CUD. The strongest signals included canonical tobacco-sensitive loci such as *AHRR* and *F2RL3*, but sensitivity analyses indicated that lifetime nicotine exposure accounted for only a limited subset of the CUD-associated DMPs. Beyond smoking-related CpGs, the methylation profile converged on immune-inflammatory, cytokine/interferon, JAK–STAT, retinoid, hemostatic, cytoskeletal, cell-adhesion and stress-regulatory pathways. Trauma burden and polysubstance-use histories explained partially distinct methylation components within the CUD group, supporting a model in which peripheral epigenetic variation in CUD reflects layered environmental and clinical exposures rather than a single drug-specific signal. Cross-tissue analyses further suggested locus-specific, but not global, convergence between blood and BA9 methylation, particularly for selected cell-migration and immune-related genes. Together, these findings nominate candidate peripheral methylation markers and immune-stress pathways for future validation, while emphasizing that CUD-associated methylation must be interpreted in the context of tobacco exposure, trauma burden, polysubstance use and tissue specificity. Longitudinal, multi-omic and multi-tissue studies will be essential to determine whether these CpGs index vulnerability, exposure-related remodeling, disease severity or treatment-relevant biological states.

Methods

Sample

Power estimations for EWAS indicate that sample sizes of approximately 400–600 participants are required to detect methylation differences of moderate magnitude at genome-wide significance thresholds [52]. Our

sample size exceeded this benchmark and provided robust statistical power consistent with previous large-scale EWAS investigations [53,54]. Participants were recruited in southern Brazil. Cases were invited from public inpatient and outpatient addiction treatment programs, and all were abstinent at the time of evaluation. The SCID-5 was used both to confirm CUD diagnosis and to assess comorbid psychiatric disorders. Controls were blood donors selected based on the absence of current substance use disorder (except tobacco), severe or unstable medical conditions, or neurological disorders, as determined by the SCID-5. All controls reported no substance use in the preceding 30 days. Informed consent was obtained from all participants in accordance with the Declaration of Helsinki and Brazilian regulations, with study approval by the Brazilian National Committee of Ethics in Research (57370016.0.1001.5336/60633516.2.1001.5327) and the Office for Human Research Protections, United States (IRB 00000921).

DNA Methylation

For each participant, a total of 10 mL of peripheral whole blood was collected by venipuncture into EDTA tubes (Becton Dickinson) and DNA isolated with the DNeasy Blood & Tissue Kit (Qiagen). The amount of DNA in each sample was quantified with a Nanodrop spectrophotometer (Thermo). Bisulfite conversion of 450 ng of DNA was performed with the EZ DNA Methylation Kit (Zymo), followed by hybridization to the Infinium EPICMethylation BeadChip Kit v.1 (Illumina), according to the manufacturer's instructions. DNA samples were systematically distributed across Illumina MethylationEPIC BeadChips to balance cases and controls within each array and minimize batch effects. Investigators were blinded to diagnostic status during preprocessing and normalization, with group labels revealed only after completion of quality control. Chips were scanned in an iScan microarray reader (Illumina). EWAS analyses were performed in R 4.4.2 using an in-house pipeline with multiple Bioconductor packages. Quality control and filtering steps were performed with *minfi* [55], including exclusion of probes with a mean detection p-value > 0.01, low bead count (< 3 in > 1% of the samples), cross-reactive probes [56], probes associated with single nucleotide polymorphisms [56], and probes located in sex chromosomes. A total of 626,937 probes remained for further analysis. Functional normalization was performed, including noob background correction, dye normalization, and correction for the first two principal components on the array [57]. Finally, probes were log-transformed from beta-values to M-values before statistical analyses [58]. Blood cell type proportions (CD8+ (CD8T) and CD4+ (CD4T) T-lymphocytes, natural killer cells (NK), B-lymphocytes (Bcell), monocytes (Mono), and neutrophils) were estimated using the *EpiDISH* package [59–64]. A DNA methylation-based estimate of smoking exposure ("smoking score") was estimated using the *EpiSmokEr* package [65,66], and we used the package EPISTRUCTURE to predict DNA methylation-based estimates of genetic ancestry from all subjects [67].

Clinical and Socio demographics variables

Sociodemographic data included age, biological sex, self-reported race/ethnicity, years of education, monthly gross income, and occupational status. Race/ethnicity was assessed using the five official categories defined by the Brazilian Institute of Geography and Statistics (IBGE): White (*Branca*), Black (*Preta*), Brown (*Parda*, referring to mixed ancestry), Yellow (*Amarela*, typically East Asian descent), and Indigenous (*Indígena*). Occupational status was categorized as employed, unemployed, student, homemaker, retired, or disabled. Marital status was classified as married/stable union, divorced/separated, single, or widowed. All categorical variables are reported as frequency (%), stratified by group (Control/CUD), as shown in **Table 1**.

Clinical variables, including the frequency of SUDs and the presence of current mood disorders, were assessed for each participant using the SCID-5. To evaluate substance use severity within the CUD group, we employed the Summary Scores for Recent Functioning (SS-Rs) derived from the Addiction Severity Index, 6th Edition (ASI-6), applying Nonparametric Item Response Theory models [68,69]. To estimate the severity of substance use in the CUD group, we calculated a lifetime regular use ratio by dividing the number of years of regular substance use by the participant's age.

Exposure to PTEs was defined as having experienced at least one event involving risk of death, violent physical harm, rape, or witnessing extreme violence. Assessment methods differed slightly between groups: the control group completed the Life Events Checklist for DSM-5 (LEC-5), whereas the CUD group was evaluated with ASI-6; [68,69]. In the ASI-6, we examined items related to rape, other violent physical assaults (e.g., beatings), life-threatening situations, and witnessing extreme violence. In the control group, we used LEC-5 items 6, 7, 8, and 11 to capture exposure to physical aggression, assault with a weapon, rape, and kidnapping, while also recording whether these events were witnessed. For both groups, we also considered severe sexual, physical, or emotional abuse in childhood, as measured by the Childhood Trauma Questionnaire (CTQ). Participants who endorsed any of the events listed above were classified as exposed to a PTE, and we created a count variable (range 0–7) reflecting the total number of PTEs reported.

Exploring the Link Between Peripheral Blood and Brain Epigenetics

This case-control study primarily aims to uncover DNA methylation alterations in peripheral blood linked to CUD. However, to deepen our understanding of CUD's pathophysiology, we initially used the BECon web application [23] to explore potential blood–brain correlations of the significant CpGs. Building on this, we expanded our analysis to include postmortem brain samples. This approach allows us to explore the connection between systemic epigenetic modifications and those within the central nervous system. Specifically, we examined whether methylation patterns identified in peripheral blood are mirrored in the dorsolateral prefrontal cortex Brodmann area (BA) 9 region of CUD subjects by assessing frozen postmortem brain samples of n=12 CUD subjects and n=16 non-psychiatric controls obtained through the UTHealth Brain Collection for Research in Psychiatric Disorders in collaboration with the Harris County Institute of Forensic Science (**Supplementary File 7**). Informed consent for the brain donation and use in this protocol was obtained from the next-of-kin. Diagnosis was confirmed through a psychological autopsy interview with the next-of-kin [70], toxicology screening, and review of relevant electronic medical records. Brain processing and dissection of the BA9 area have been recently described [71]. DNA was isolated using the DNeasy Blood & Tissue kit, and DNA methylation levels were assessed following the same methods described above, in blood.

Statistical Analyses

All statistical analyses comparing groups were conducted using two-sided tests unless otherwise specified. Differentially methylated positions (DMPs) between cases and controls were identified using a linear mixed-effects model, with DNA methylation levels as the dependent variable, chip (Sentrix_ID) as a random-effects term, and chip position (Sentrix_Position), age, sex, CD4T, CD8T, NK cells, B cells, monocytes, and the first 10 EPISTRUCTURE principal components (to adjust for unmeasured ancestry-related variation in methylation profiles) as fixed-effects covariates. To further address potential confounding by population stratification in DNA methylation, we systematically evaluated whether our top DMPs overlapped with CpG sites previously reported to capture ancestry-related signal. For this purpose, we used the catalog from Galanter et al. [72] which identified 916 CpG sites associated with ethnicity in an admixed Latino cohort, including both ancestry-independent and ancestry-driven subsets. This step served as an additional safeguard against ancestry-driven bias, given that our two study populations differed substantially in self-reported skin color.

To evaluate whether CUD-associated methylation differences could be explained by nicotine exposure, we performed a sensitivity EWAS restricted to individuals with CUD. DNA methylation M-values were modeled as a function of lifetime nicotine exposure, defined as years of regular nicotine use divided by age. The model included the same covariate structure used in the primary EWAS, with Sentrix ID modeled as a random intercept and Sentrix position, age, sex, estimated blood cell proportions, and the first ten EPISTRUCTURE principal components included as fixed-effect covariates. CpGs associated with nicotine exposure were defined using the same exploratory significance threshold applied to the primary top-ranked DMP analysis, namely FDR-adjusted $p < 5 \times 10^{-3}$. Nicotine-associated CpGs were then compared with the CUD-associated DMPs identified in the primary case-control EWAS. Overlapping CpGs were classified as directionally concordant when the sign of the nicotine effect matched the sign of the CUD-control effect.

Scripts for the EWAS analyses were adapted from those shared by the PGC PTSD EWAS Working Group [73] Correction for multiple testing was performed using the Benjamini–Hochberg false discovery rate (FDR) procedure. Following established practices in EWAS [74] we adopted a conservative genome-wide significance Bonferroni threshold of $p < 9 \times 10^{-8}$, which accounts for the effective number of independent CpGs interrogated. However, Bonferroni correction is often regarded as overly conservative, as DNA methylation levels are correlated across neighboring CpG sites and many probes on the array are non-variable [75] , leading to an overestimation of the number of truly independent tests. To address this limitation and provide a more interpretable set of results within the constraints of our sample size, we additionally report top-ranked DMPs defined by an FDR-adjusted $p < 5 \times 10^{-3}$. This dual-threshold strategy, determined *a priori*, enables us to identify both the most statistically robust associations and biologically plausible signals that, while not passing the strictest genome-wide correction, may merit further investigation in larger or independent samples. To evaluate and adjust for potential bias and inflation of test statistics, we calculated the genomic inflation factor (λ), generated quantile–quantile (Q–Q) plots, and applied Bayesian adjustment using the R package BACON [76].

To examine associations between lifetime exposure to PTEs, SUD severity, socioeconomic variability, and differentially methylated CpG sites, we first conducted a principal component analysis (PCA) to reduce the dimensionality of significant CpG sites identified in the EWAS. All significantly differentially methylated CpGs were entered into the PCA, and the rotated solution obtained with Promax rotation and Kaiser normalization yielded principal components (PCs) with eigenvalues greater than 1. Factor loadings greater than 0.40 were considered meaningful. The Kaiser–Meyer–Olkin (KMO) measure was calculated to confirm sampling adequacy. Individual PC scores were then retained for each participant within the CUD group. These PC scores served as dependent variables in stepwise linear regressions, including as candidate predictors: (i) number of lifetime PTEs, (ii) ratio of years of regular cannabis, alcohol, nicotine, or cocaine use to age, and (iii) monthly income as a socioeconomic indicator. The most parsimonious statistically significant models were retained, and semi-partial R^2 values of significant predictors were calculated to quantify their relative contributions.

Differences between the CUD and control groups in postmortem BA9 samples were assessed by linear regression, using group membership as the independent variable and methylation levels of previously identified blood-derived significant CpGs as dependent variables, with age, sex, smoking score, postmortem interval (PMI), and brain tissue pH as covariates. A significance threshold of FDR-adjusted p-value < 0.05 was applied.

Bioinformatic Annotation and Functional Enrichment Analysis

All assayed CpGs were annotated for gene associations and genomic context using the official Infinium MethylationEPIC annotation package (IlluminaHumanMethylationEPICanno.ilm10b4.hg19) in R version 4.5.0. This annotation was merged with the EWAS results matrix to assign gene overlaps, CpG island context, and promoter-proximity annotations. CpGs lacking valid annotations were excluded from downstream enrichment analyses. Top-ranked DMPs were defined as CpGs with FDR-adjusted $p < 5 \times 10^{-3}$ and were stratified according to direction of effect, distinguishing CpGs with lower methylation in CUD from those with higher methylation in CUD. When individual CpGs mapped to multiple genes, gene assignments were split into unique gene identifiers. Enrichment or depletion of top-ranked DMPs across CpG island contexts was assessed by comparing their observed distribution with the distribution of all retained EPIC probes after quality control. Odds ratios and log2-transformed odds ratios were used to quantify and visualize overrepresentation or depletion across CpG island annotation categories.

For KEGG pathway enrichment, gene symbols annotated to direction-stratified DMPs were converted to Entrez IDs using the bitr() function from clusterProfiler [77], referencing the org.Hs.eg.db database. KEGG enrichment was performed using enrichKEGG() in clusterProfiler, separately for genes annotated to hypomethylated and hypermethylated DMPs. Statistical significance was defined as FDR-adjusted $p < 0.05$, and results were visualized using custom ggplot2 dotplots [78].

As a complementary exploratory analysis, the same direction-stratified gene lists were submitted to Enrichr (release 20 March 2025) to query additional pathway, disease/trait, transcription-factor, and perturbation databases. The queried libraries included Gene Ontology, Reactome 2024, BioPlanet 2019, MSigDB Hallmark 2020, GWAS Catalog 2023, UK Biobank GWAS v1, DisGeNET 7.0, Orphanet Augmented 2021, ClinVar 2019, ChEA 2022, HDSigDB 2021, Virus Perturbations from GEO, and VirusMINT. Enrichr performs two-tailed Fisher exact tests against each library-specific background universe; q-values < 0.05 after Benjamini–Hochberg correction within each library were considered significant. Results were ordered by Enrichr's combined score, calculated from the log-transformed p-value and the z-score deviation from the expected rank. Redundant terms with greater than 50% gene overlap were collapsed to the highest-scoring parent term to reduce inflation of related categories.

Protein–protein interaction analyses were performed separately for genes annotated to CpGs with lower and higher methylation in CUD using STRING v12.0 [79]. Submitted gene sets were analyzed using a medium interaction confidence threshold of 0.4. Only interactions supported by curated databases, experimental evidence, or co-expression were retained. No additional interactors were added beyond the submitted genes. STRING's permutation framework compares the observed number of edges with random networks of identical size; PPI enrichment was reported as a one-sided p-value.

Data availability

Due to ethical and legal restrictions under Brazilian regulations governing the sharing of clinical and personally identifiable information, individual-level clinical data cannot be made publicly available. Access to de-identified participant-level data for scientifically valid research purposes can be granted upon reasonable request, subject to approval by the institutional ethics committee and the data governance board of the Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Requests should be directed to the corresponding author, who will facilitate the review and authorization process in compliance with all applicable ethical and legal requirements. All biological data underlying the EWAS analyses, including methylation beta values, differential methylation statistics, and related summary data, are fully available within the manuscript and its Supplementary Information.

References

1. Schwartz, E. K. C., Wolkowicz, N. R., De Aquino, J. P., MacLean, R. R. & Sofuoglu, M. Cocaine use disorder (CUD): Current clinical perspectives. *Subst. Abuse Rehabil.* **13**, 25–46 (2022).
2. United Nations Office on Drugs and Crime. *World Drug Report 2024*. (United Nations, 2024). doi:10.18356/9789211065534.
3. Tondo, L. P. *et al.* White matter deficits in cocaine use disorder: convergent evidence from in vivo diffusion tensor imaging and ex vivo proteomic analysis. *Transl. Psychiatry* **11**, 252 (2021).
4. Rothmann, L. M. *et al.* The cortical thickness of tricenarian cocaine users assembles features of an octogenarian brain. *J. Neurosci. Res.* **102**, e25287 (2024).
5. Sun, J., Kranzler, H. R., Gelernter, J. & Bi, J. A genome-wide association study of cocaine use disorder accounting for phenotypic heterogeneity and gene–environment interaction. *J. Psychiatry Neurosci.* **45**,

34–44 (2020).

6. Cabana-Domínguez, J., Shivalikanjli, A., Fernàndez-Castillo, N. & Cormand, B. Genome-wide association meta-analysis of cocaine dependence: Shared genetics with comorbid conditions. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **94**, 109667 (2019).
7. Fernàndez-Castillo, N., Cabana-Domínguez, J., Corominas, R. & Cormand, B. Molecular genetics of cocaine use disorders in humans. *Mol. Psychiatry* **27**, 624–639 (2022).
8. Massart, R. *et al.* Role of DNA methylation in the nucleus accumbens in incubation of cocaine craving. *J. Neurosci.* **35**, 8042–8058 (2015).
9. Zillich, E. *et al.* Multi-omics profiling of DNA methylation and gene expression alterations in human cocaine use disorder. *medRxiv* (2024) doi:10.1101/2024.04.24.24306302.
10. Camilo, C. *et al.* Genome-wide DNA methylation profile in the peripheral blood of cocaine and crack dependents. *Braz J Psychiatry* **41**, 485–493 (2019).
11. Shu, C. *et al.* Epigenome-wide association analyses of active injection drug use. *Drug Alcohol Depend.* **235**, 109431 (2022).
12. Viola, T. W. *et al.* Prenatal cocaine exposure and its influence on pediatric epigenetic clocks and epigenetic scores in humans. *Sci. Rep.* **14**, 1946 (2024).
13. Hughes, K., Ford, K., Kadel, R., Sharp, C. A. & Bellis, M. A. Health and financial burden of adverse childhood experiences in England and Wales: a combined primary data study of five surveys. *BMJ Open* **10**, e036374 (2020).
14. Carliner, H. *et al.* Childhood trauma and illicit drug use in adolescence: A population-based national comorbidity survey replication–adolescent supplement study. *J. Am. Acad. Child Adolesc. Psychiatry* **55**, 701–708 (2016).
15. Afful, S. E., Strickland, J. R., Cottler, L. & Bierut, L. J. Exposure to trauma: a comparison of cocaine-dependent cases and a community-matched sample. *Drug Alcohol Depend.* **112**, 46–53 (2010).
16. Sanvicente-Vieira, B. *et al.* Sex-based differences in multidimensional clinical assessments of early-abstinence crack cocaine users. *PLoS One* **14**, e0218334 (2019).

17. Roglio, V. S. *et al.* Prediction of attempted suicide in men and women with crack-cocaine use disorder in Brazil. *PLoS One* **15**, e0232242 (2020).
18. Bruxel, E. M. *et al.* Psychiatric genetics in the diverse landscape of Latin American populations. *Nat. Genet.* **57**, 1074–1088 (2025).
19. Fasanelli, F. *et al.* Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts. *Nat. Commun.* **6**, 10192 (2015).
20. Andersen, A. M., Dogan, M. V., Beach, S. R. H. & Philibert, R. A. Current and future prospects for epigenetic biomarkers of substance use disorders. *Genes (Basel)* **6**, 991–1022 (2015).
21. *Epigenetic Changes of FKBP5 as a Link Connecting Stress and Psychiatric Disorder*.
22. Wielscher, M. *et al.* DNA methylation signature of chronic low-grade inflammation and its role in cardio-respiratory diseases. *Nat. Commun.* **13**, 2408 (2022).
23. Edgar, R. D., Jones, M. J., Meaney, M. J., Turecki, G. & Kobor, M. S. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl. Psychiatry* **7**, e1187 (2017).
24. Kalman, D., Morissette, S. B. & George, T. P. Co-morbidity of smoking in patients with psychiatric and substance use disorders. *Am. J. Addict.* **14**, 106–123 (2005).
25. Roll, J. M., Higgins, S. T., Budney, A. J., Bickel, W. K. & Badger, G. J. A comparison of cocaine-dependent cigarette smokers and non-smokers on demographic, drug use and other characteristics. *Drug Alcohol Depend.* **40**, 195–201 (1996).
26. de Vries, M. *et al.* From blood to lung tissue: effect of cigarette smoke on DNA methylation and lung function. *Respir. Res.* **19**, 212 (2018).
27. Bravo-Gutiérrez, O. A. *et al.* Hypomethylation of AHRR (cg05575921) Is Related to Smoking Status in the Mexican Mestizo Population. *Genes* **12**, (2021).
28. Grieshober, L. *et al.* AHRR methylation in heavy smokers: associations with smoking, lung cancer risk, and lung cancer mortality. *BMC Cancer* **20**, 905 (2020).
29. López, A. J., Siciliano, C. A. & Calipari, E. S. Activity-Dependent Epigenetic Remodeling in Cocaine Use Disorder. *Handb. Exp. Pharmacol.* **258**, 231–263 (2020).

30. Skov-Jeppesen, S. M., Kobylecki, C. J., Jacobsen, K. K. & Bojesen, S. E. Changing Smoking Behavior and Epigenetics: A Longitudinal Study of 4,432 Individuals From the General Population. *Chest* **163**, 1565–1575 (2023).
31. Bojesen, S. E., Timpson, N., Relton, C., Davey Smith, G. & Nordestgaard, B. G. AHRR (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax* **72**, 646–653 (2017).
32. Smith, A. K. *et al.* Epigenome-wide meta-analysis of PTSD across 10 military and civilian cohorts identifies methylation changes in AHRR. *Nat. Commun.* **11**, 5965 (2020).
33. Riaz, F., Pan, F. & Wei, P. Aryl hydrocarbon receptor: The master regulator of immune responses in allergic diseases. *Front. Immunol.* **13**, 1057555 (2022).
34. Gutiérrez-Vázquez, C. & Quintana, F. J. Regulation of the Immune Response by the Aryl Hydrocarbon Receptor. *Immunity* **48**, 19–33 (2018).
35. Stamatovich, S. N. *et al.* Plasma pro- and anti-inflammatory cytokines may relate to cocaine use, cognitive functioning, and depressive symptoms in cocaine use disorder. *Am. J. Drug Alcohol Abuse* **47**, 52–64 (2021).
36. Moreira, F. P. *et al.* Cocaine abuse and effects in the serum levels of cytokines IL-6 and IL-10. *Drug Alcohol Depend.* **158**, 181–185 (2016).
37. Ersche, K. D. & Döffinger, R. Inflammation and infection in human cocaine addiction. *Current Opinion in Behavioral Sciences* **13**, 203–209 (2017).
38. Carow, B. & Rottenberg, M. E. SOCS3, a Major Regulator of Infection and Inflammation. *Front. Immunol.* **5**, 58 (2014).
39. Kobayashi, K. S. & van den Elsen, P. J. NLRC5: a key regulator of MHC class I-dependent immune responses. *Nat. Rev. Immunol.* **12**, 813–820 (2012).
40. A, Z., Cd, K., Ad, V., T, K. & Li, S. Expression of Retinoic Acid Receptor (RAR) α Protein in the Synovial Membrane from Patients with Osteoarthritis and Rheumatoid Arthritis. *Int. J. Biomed. Sci.* **3**, 46–49 (2007).
41. Connelly, K. L. & Unterwald, E. M. Chronic cocaine administration upregulates FKBP5 in the extended

amygdala of male and female rats. *Drug Alcohol Depend.* **199**, 101–105 (2019).

42. Yu, H. *et al.* Epigenome-wide association study identifies Behçet's disease-associated methylation loci in Han Chinese. *Rheumatology* **58**, 1574–1584 (2019).
43. Fries, G. R., Gassen, N. C. & Rein, T. The FKBP51 Glucocorticoid Receptor Co-Chaperone: Regulation, Function, and Implications in Health and Disease. *Int. J. Mol. Sci.* **18**, (2017).
44. Klengel, T. *et al.* Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat. Neurosci.* **16**, 33–41 (2013).
45. Fox, H. C. *et al.* Immune system inflammation in cocaine dependent individuals: implications for medications development. *Hum. Psychopharmacol.* **27**, 156–166 (2012).
46. Restrepo, C. S. *et al.* Pulmonary complications from cocaine and cocaine-based substances: imaging manifestations. *Radiographics* **27**, 941–956 (2007).
47. Zaparte, A. *et al.* Cocaine use disorder is associated with changes in Th1/Th2/Th17 cytokines and lymphocytes subsets. *Front. Immunol.* **10**, 2435 (2019).
48. Xie, J. & Wang, H. Semaphorin 7A as a potential immune regulator and promising therapeutic target in rheumatoid arthritis. *Arthritis Res. Ther.* **19**, 10 (2017).
49. Inoue, N., Nishizumi, H., Naritsuka, H., Kiyonari, H. & Sakano, H. Sema7A/PlxnCl signaling triggers activity-dependent olfactory synapse formation. *Nat. Commun.* **9**, 1842 (2018).
50. Wang, J. & Deretic, D. The Arf and Rab11 effector FIP3 acts synergistically with ASAP1 to direct Rabin8 in ciliary receptor targeting. *J. Cell Sci.* **128**, 1375–1385 (2015).
51. Villicaña, S. & Bell, J. T. Genetic impacts on DNA methylation: research findings and future perspectives. *Genome Biol.* **22**, 127 (2021).
52. Tsai, P.-C. & Bell, J. T. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *Int. J. Epidemiol.* **44**, 1429–1441 (2015).
53. Zillich, L. *et al.* Epigenome-wide association study of alcohol use disorder in five brain regions. *Neuropsychopharmacology* **47**, 832–839 (2022).
54. Hannon, E. *et al.* An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization

of genetic associations and differential DNA methylation. *Genome Biol.* **17**, 176 (2016).

55. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
56. McCartney, D. L. *et al.* Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genom Data* **9**, 22–24 (2016).
57. Fortin, J.-P. *et al.* Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).
58. Du, P. *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* **11**, 587 (2010).
59. Zheng, S. C. *et al.* A novel cell-type deconvolution algorithm reveals substantial contamination by immune cells in saliva, buccal and cervix. *Epigenomics* **10**, 925–940 (2018).
60. Zheng, S. C., Breeze, C. E., Beck, S. & Teschendorff, A. E. Identification of differentially methylated cell types in epigenome-wide association studies. *Nat. Methods* **15**, 1059–1066 (2018).
61. Teschendorff, A. E., Breeze, C. E., Zheng, S. C. & Beck, S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinformatics* **18**, 105 (2017).
62. Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
63. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).
64. Teschendorff, A. E. & Zheng, S. C. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics* **9**, 757–768 (2017).
65. Bollepalli, S., Korhonen, T., Kaprio, J., Anders, S. & Ollikainen, M. EpiSmokEr: a robust classifier to determine smoking status from DNA methylation data. *Epigenomics* **11**, 1469–1486 (2019).
66. Elliott, H. R. *et al.* Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clin. Epigenetics* **6**, 4 (2014).

67. Rahmani, E. *et al.* Genome-wide methylation data mirror ancestry information. *Epigenetics Chromatin* **10**, 1 (2017).

68. Denis, C. M., Cacciola, J. S. & Alterman, A. I. Addiction Severity Index (ASI) summary scores: comparison of the Recent Status Scores of the ASI-6 and the Composite Scores of the ASI-5. *J. Subst. Abuse Treat.* **45**, 444–450 (2013).

69. Kessler, F. *et al.* Psychometric properties of the sixth version of the Addiction Severity Index (ASI-6) in Brazil. *Braz J Psychiatry* **34**, 24–33 (2012).

70. Meyer, T. D., Godfrey, C. J. & Walss-Bass, C. The UT health Psychological Autopsy Interview Schedule (UTH- PAIS) - Description and reliability of diagnoses and transdiagnostic personality measures. *J. Psychiatr. Res.* **156**, 221–227 (2022).

71. Cabrera-Mendoza, B. *et al.* Within subject cross-tissue analyzes of epigenetic clocks in substance use disorder postmortem brain and blood. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **192**, 13–27 (2023).

72. Galanter, J. M. *et al.* Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *Elife* **6**, (2017).

73. Katrinli, S. *et al.* Epigenome-wide association studies identify novel DNA methylation sites associated with PTSD: a meta-analysis of 23 military and civilian cohorts. *Genome Med.* **16**, 147 (2024).

74. Mansell, G. *et al.* Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. *BMC Genomics* **20**, 366 (2019).

75. Ong, M.-L. & Holbrook, J. D. Novel region discovery method for Infinium 450K DNA methylation data reveals changes associated with aging in muscle and neuronal pathways. *Aging Cell* **13**, 142–155 (2014).

76. van Iterson, M., van Zwet, E. W., BIOS Consortium & Heijmans, B. T. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol.* **18**, 19 (2017).

77. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

78. Villanueva, R. A. M. & Chen, Z. J. ggplot2: Elegant Graphics for Data Analysis (2nd ed.). *Measurement*

(Mahwah NJ) **17**, 160–167 (2019).

79. von Mering, C. *et al.* STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* **33**, D433-7 (2005).

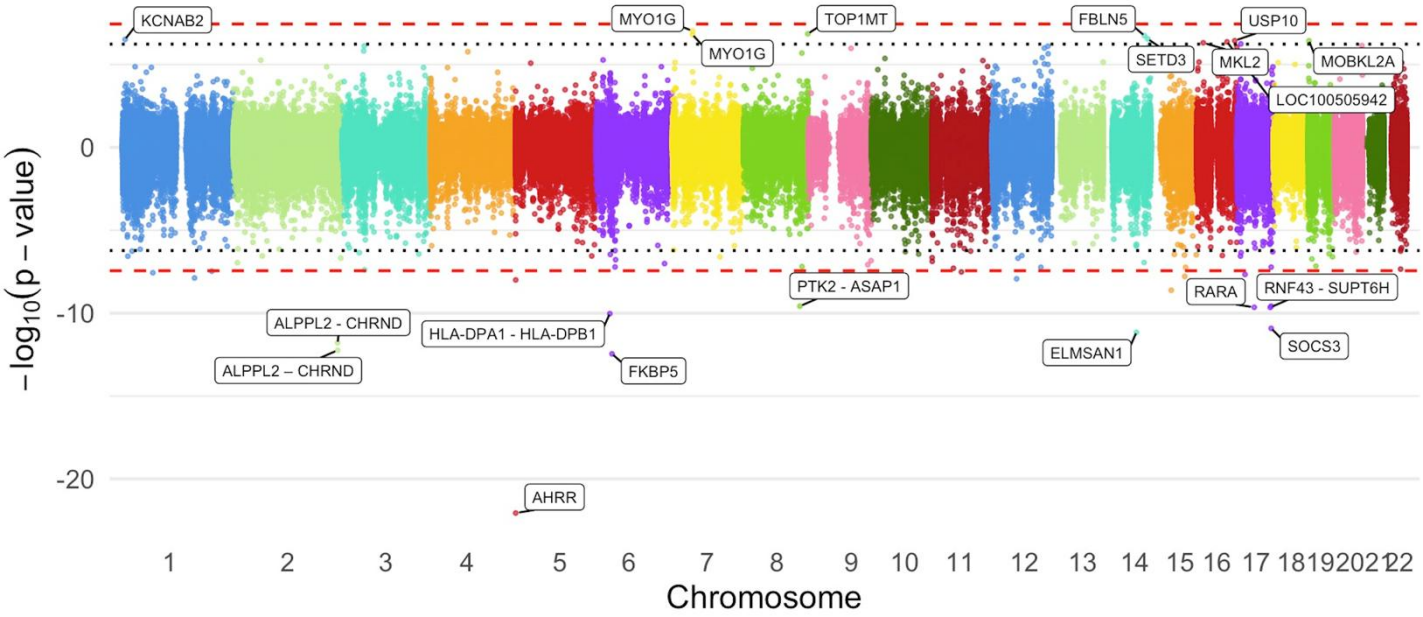


Figure 1. Miami plot. The x-axis is the chromosomal location of each site across the genome. The y-axis is the −log10 of the unadjusted p-value for the association with cocaine use disorder. The line (black) represents the epigenome-wide significance threshold of FDR<.005 and red line Bonferroni significance threshold of p< 3.6e-08.

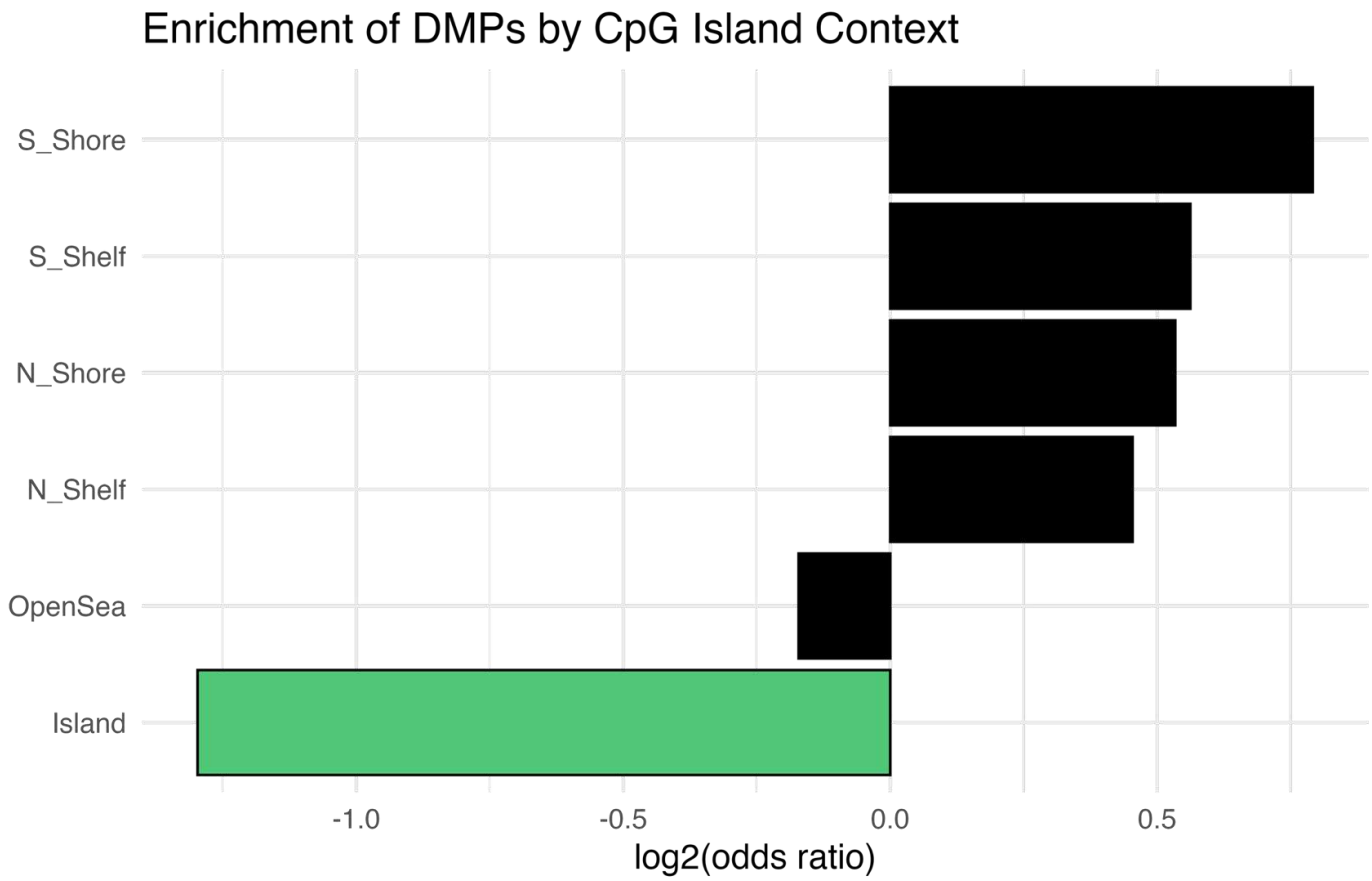


Figure 2. Relative enrichment of CUD-associated DMPs across CpG island contexts.

Bar plot showing log2 odds ratios for the enrichment or depletion of differentially methylated positions across CpG island annotation categories relative to the EPIC array background. Positive values indicate overrepresentation among DMPs, whereas negative values indicate depletion. Although OpenSea regions accounted for the largest fraction of DMPs in absolute annotation counts, they were slightly depleted after normalization to the array background. In contrast, DMPs were overrepresented in CpG island flanking regions, particularly S_Shore and S_Shelf, and depleted in CpG islands. This pattern indicates that CUD-associated methylation differences preferentially map to CpG island-adjacent regulatory contexts when accounting for the genomic distribution of interrogated CpGs.

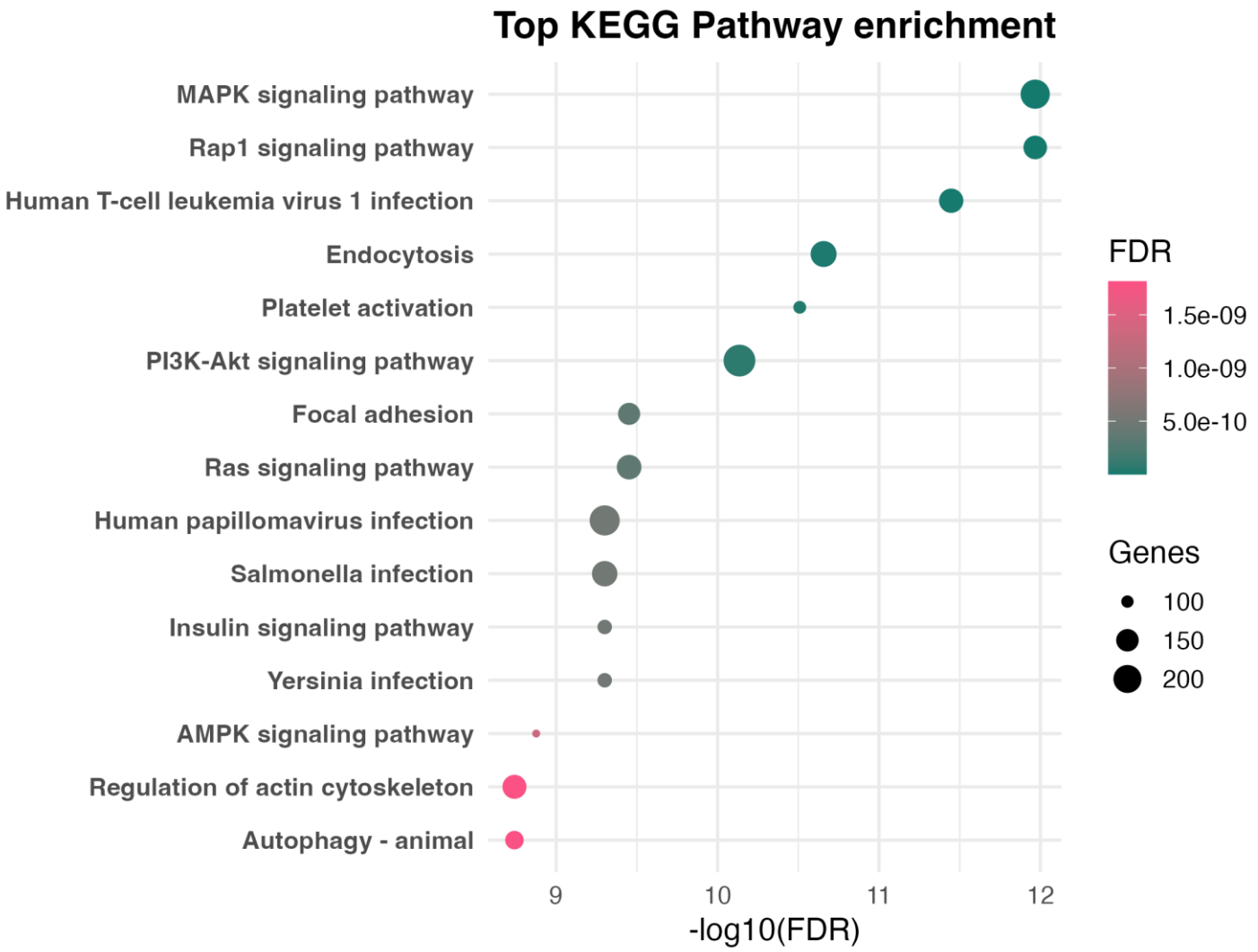


Figure 3. Top KEGG pathways enriched among genes annotated to CUD-associated DMPs.

Dot plot showing the top KEGG pathways enriched among genes annotated to differentially hypomethylated positions associated with cocaine use disorder. The x-axis represents enrichment significance as −log10(FDR). Dot size indicates the number of genes contributing to each pathway, and dot color represents the FDR-adjusted significance. Overall, the enriched pathways converge on intracellular signal transduction, immune-inflammatory activation, cytoskeletal remodeling, adhesion, endocytosis, and metabolic stress-response mechanisms.

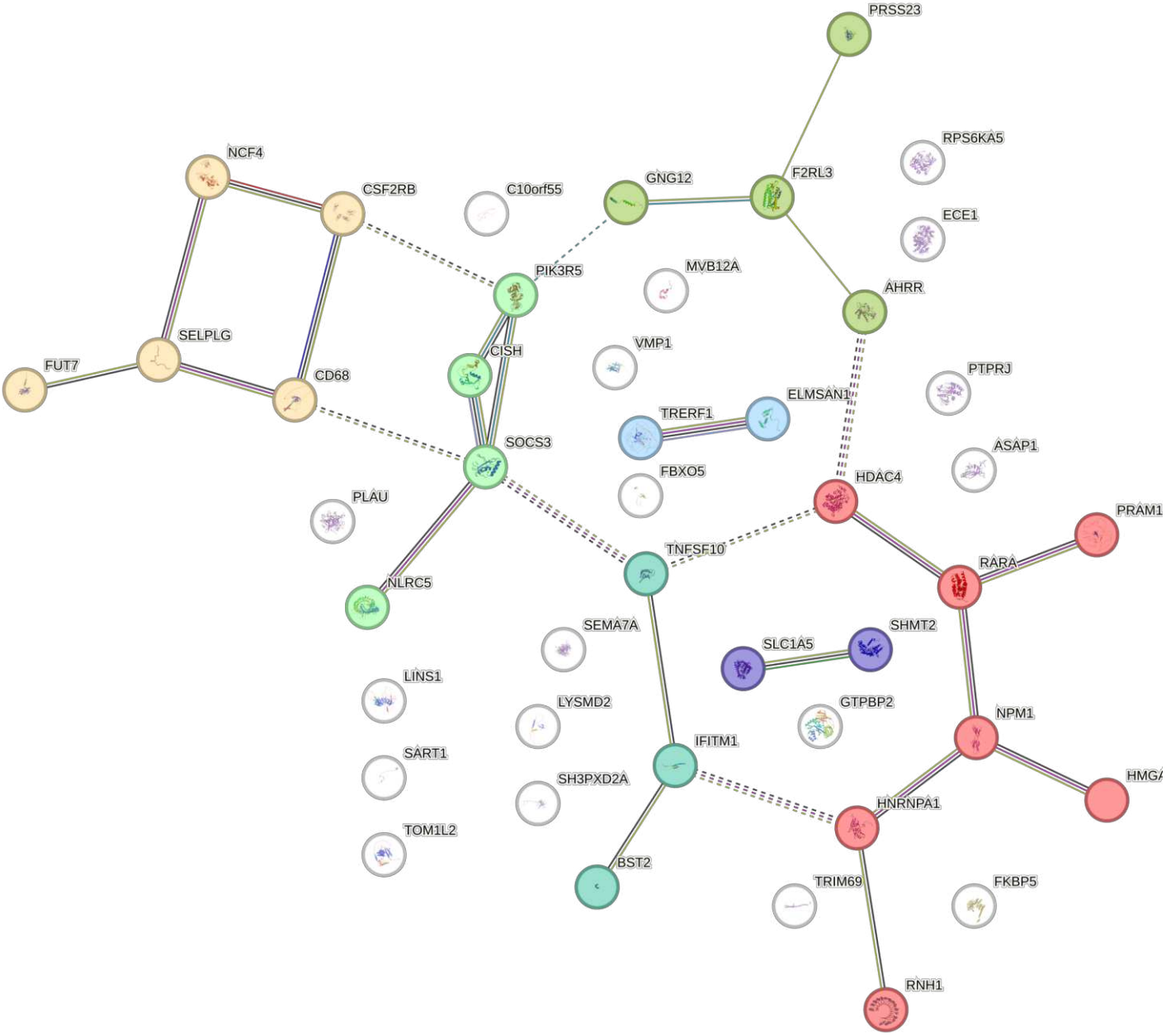


Figure 4. Protein–protein interaction network of genes annotated to hypomethylated DMPs in CUD.

Genes annotated to hypomethylated differentially methylated positions were submitted to STRING v12 using a minimum interaction confidence score of 0.4. The network comprised 45 nodes connected by 29 edges, exceeding the 13 edges expected by chance and indicating significant protein–protein interaction enrichment (PPI enrichment $p = 7.98 \times 10^{-5}$). Nodes represent proteins encoded by genes annotated to hypomethylated CpG sites, and edges represent known or predicted functional associations. The network highlights functionally connected subnetworks involving innate immune regulation, cytokine/JAK–STAT signaling, interferon-associated responses, leukocyte adhesion, cytoskeletal remodeling, and transcriptional/chromatin regulation. Prominent connected genes include SOCS3, CISH, TNFSF10, IFITM1, BST2, CSF2RB, CD68, SELPLG, NCF4, and NLRC5, consistent with immune-inflammatory pathway enrichment. Additional connected or peripheral nodes, including AHRR, F2RL3, HDAC4, RARA, NPM1, HNRNPA1, and FKBP5, suggest convergence between environmental exposure-sensitive loci, transcriptional regulation, and immune signaling. Overall, the excess connectivity supports that CUD-associated hypomethylation reflects a biologically organized inflammatory and regulatory signature rather than isolated epigenetic variation.

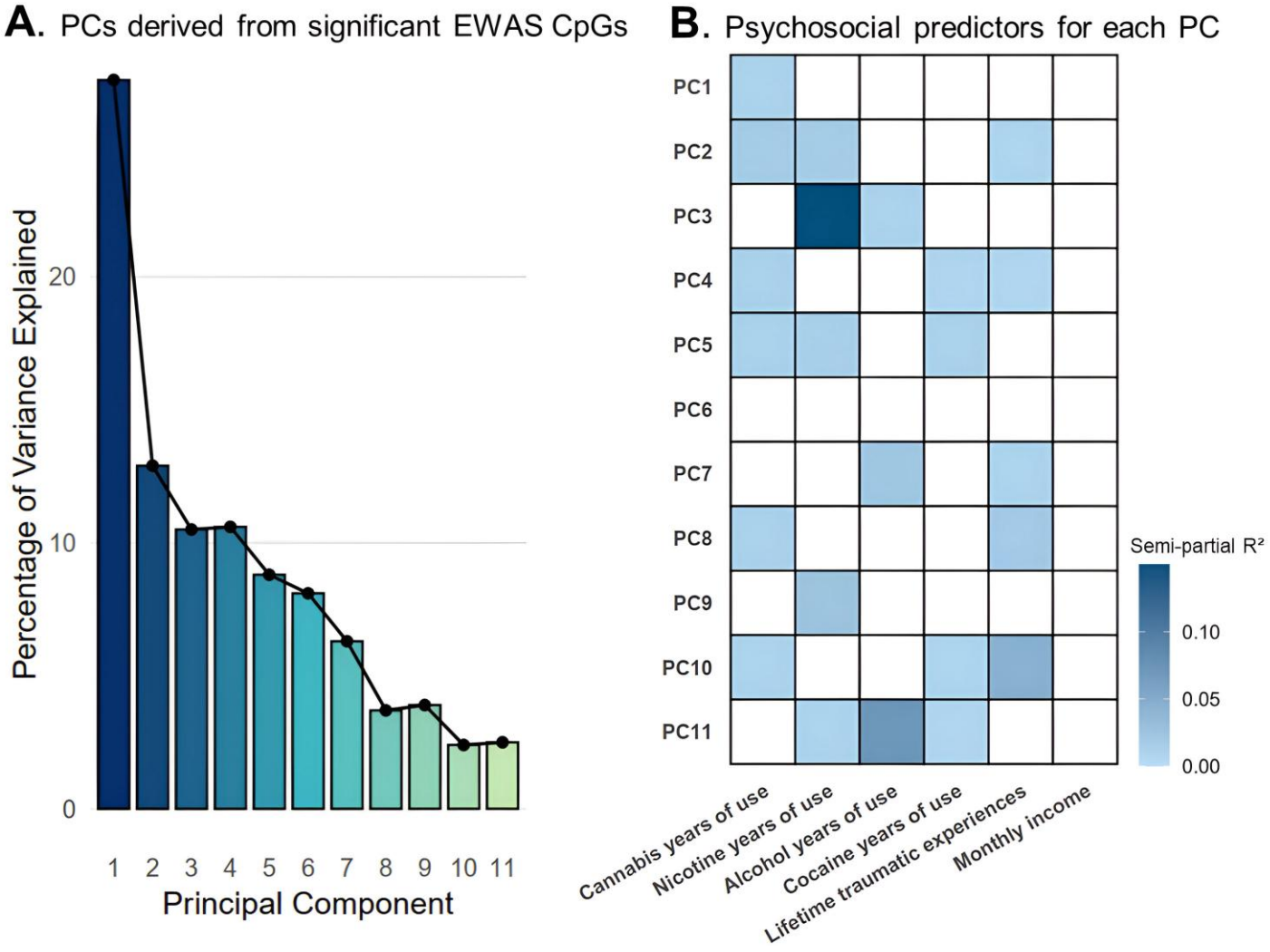


Figure 5. Principal component structure of EWAS-significant CpG sites and psychosocial predictors of epigenetic variation within the CUD group.

A. Principal component analysis was performed using the 79 CpG sites identified as significantly differentially methylated in the EWAS. The rotated solution yielded 11 principal components with eigenvalues greater than 1, cumulatively explaining 76.3% of the methylation variance. Bars indicate the percentage of variance explained by each component, and the overlaid line shows the corresponding scree profile.

B. Heatmap showing significant psychosocial and substance-use predictors of each methylation-derived principal component among individuals with cocaine use disorder. Candidate predictors included lifetime cannabis, nicotine, alcohol, and cocaine exposure, expressed as years of regular use relative to age, number of lifetime potentially traumatic events, and monthly income. Color intensity represents the semi-partial R^2 of

significant predictors retained in the final regression models, reflecting their unique contribution to each component after accounting for the other predictors. White cells indicate non-significant predictors. Trauma burden and polysubstance exposure, but not monthly income, explained partially distinct patterns of methylation variance across components.

Table 1. Sample characteristics.

	Valid entries (Control/CUD)	CUD (n=563) mean (SD)	CUD median [IQR]	Control (n=687) mean (SD)	Control median [IQR]	p-value
Age	(683/562)	34.96 (9.75)	34.00 [27.25, 41.00]	34.13 (11.36)	32.00 [24.00, 42.00]	0.017
Sex: Women, freq. (%)	(683/562)	345 (61.4)		370 (54.2)		0.012
BMI	(661/436)	23.50 (4.04)	22.93 [21.00, 25.39]	26.17 (4.46)	25.46 [22.99, 28.63]	<0.001
Monthly gross Income (Brazilian Real)	(528/562)	606.78 (3105.03)	0.00 [0.00, 587.50]	6049.10 (5084.50)	4406.83 [2937.89, 7344.72]	<0.001
Years of school education	(663/538)	7.07 (3.28)	6.50 [6.50, 9.50]	14.57 (2.97)	15.00 [12.00, 17.00]	<0.001
Race/ethnicity						
White, freq. (%)	(682/533)	232 (43.5)		559 (82.0)		<0.001
Black, freq. (%)	(682/533)	206 (38.6)		46 (6.7)		<0.001
Mixed, freq. (%)	(682/533)	89 (16.7)		70 (10.3)		0.001
Indigenous, freq. (%)	(682/533)	5 (0.9)		5 (0.7)		0.942
Asian, freq. (%)	(682/533)	1 (0.2)		2 (0.3)		1
Occupation						
Employed, freq. (%)	(663/351)	73 (20.8)		480 (72.4)		<0.001
Unemployed, freq. (%)	(663/351)	149 (42.5)		22 (3.3)		<0.001
Student, freq. (%)	(663/351)	3 (0.9)		130 (19.6)		<0.001
HousewifeHome, freq. (%)	(663/351)	42 (12.0)		14 (2.1)		<0.001
Retired, freq. (%)	(663/351)	13 (3.7)		17 (2.6)		0.41
Disable, freq. (%)	(663/351)	33 (9.4)		0 (0.0)		<0.001
MaritalStatus						
Married/Stable Union, freq. (%)	(661/535)	160 (29.9)		250 (37.8)		0.005
Divorced/Separated, freq. (%)	(661/535)	160 (29.9)		29 (4.4)		<0.001
Single, freq. (%)	(661/535)	172 (32.1)		379 (57.3)		<0.001
Widow, freq. (%)	(661/535)	20 (3.7)		3 (0.5)		<0.001
Lifetime Substance Use Disorder						
Nicotine, freq. (%)	(683/562)	495 (88.1)		129 (18.9)		<0.001
Smoking Score[1]	(679/562)	5.99 (3.86)	5.97 [3.36, 8.47]	0.53 (2.89)	0.15 [-1.35, 1.77]	<0.001
Alcohol, freq. (%)	(683/562)	386 (68.7)		8 (1.2)		<0.001
Cannabis, freq. (%)	(683/562)	412 (73.3)		3 (0.4)		<0.001
Amphetamines, freq. (%)	(683/562)	31 (5.5)		0 (0.0)		<0.001
Opioids, freq. (%)	(683/562)	8 (1.4)		0 (0.0)		0.006
Sedatives, freq. (%)	(683/562)	84 (14.9)		0 (0.0)		<0.001
Psychiatric Comorbidity						
Bipolar Disorder, freq. (%)	(662/408)	30 (7.4)		25 (3.8)		0.015

Substance-Induced Bipolar Disorder, freq. (%)	(662/408)	33 (8.1)	0 (0.0)	<0.001
Depressive Disorder, freq. (%)	(662/408)	136 (33.3)	152 (23.0)	<0.001
Substance-Induced Depressive Disorder, freq. (%)	(662/408)	62 (15.2)	0 (0.0)	<0.001

Note. CUD, Cocaine Use Disorder; differences between continuous variables were tested using the Mann-Whitney test, which considered the non-normal distribution of the data, as determined by the Shapiro-Wilk test; differences in frequencies were measured using chi-square tests. [1] DNA methylation-based estimate of smoking exposure estimated using the EpiSmokEr.

Table 2. Top Differentially Methylated Positions (DMPs) Associated with CUD

CpG	Effect Size	SE	t Statistics	p value	FDR	Gene	Methylation
cg05575921	-1.1774273	0.11982322	-9.8263702	8.67E-23	5.44E-17	AHRR	Reduced
cg03546163	-0.5429421	0.07464095	-7.2740521	3.49E-13	1.09E-07	FKBP5	Reduced
cg01940273	-0.3394011	0.04709731	-7.2063804	5.75E-13	1.20E-07	AC068134.5	Reduced
cg21566642	-0.4789623	0.06781946	-7.0623131	1.64E-12	2.57E-07	AC068134.5	Reduced
cg02738868	-0.1821331	0.02655128	-6.8596742	6.90E-12	8.66E-07	ELMSAN1	Reduced
cg13343932	-0.4592058	0.06772751	-6.7801966	1.20E-11	1.25E-06	SOCS3	Reduced
cg14753356	-0.3798091	0.05863942	-6.4770264	9.35E-11	8.38E-06	Intergenic	Reduced
cg11047325	-0.5360339	0.08483632	-6.3184488	2.64E-10	1.51E-05		Reduced
cg13117318	-0.3889299	0.06129285	-6.345436	2.22E-10	1.51E-05	RARA	Reduced
cg17739917	-0.3949416	0.06226055	-6.3433691	2.25E-10	1.51E-05	Intergenic	Reduced
cg26140475	-0.3180252	0.05032354	-6.3196104	2.62E-10	1.51E-05	SOCS3	Reduced
cg03636183	-0.3399521	0.05551088	-6.1240635	9.12E-10	4.77E-05	F2RL3	Reduced
cg18181703	-0.2690717	0.04445533	-6.0526313	1.42E-09	6.87E-05	SOCS3	Reduced
cg15022400	-0.3255845	0.05454743	-5.9688322	2.39E-09	0.00010704	TRIM69	Reduced
cg08913523	-0.3432329	0.0588585	-5.8314922	5.49E-09	0.00022967	Intergenic	Reduced
cg12054453	-0.6021715	0.10493088	-5.7387447	9.54E-09	0.00037023	TMEM49	Reduced
cg21161138	-0.3171049	0.05534002	-5.7301188	1.00E-08	0.00037023	AHRR	Reduced
cg02583484	-0.231711	0.04062638	-5.7034623	1.17E-08	0.00040902	HNRNPA1;HNRPA1L-2	Reduced
cg11320225	-0.2225869	0.03916164	-5.6837977	1.32E-08	0.00043482	Intergenic	Reduced
cg18110140	-0.2946001	0.05217837	-5.6460197	1.64E-08	0.00051489	PPCDC	Reduced
cg00303773	-0.4179135	0.07464674	-5.5985501	2.16E-08	0.0006455	TOM1L2	Reduced
cg25189904	-0.3825894	0.06884213	-5.5574902	2.74E-08	0.00078015	GNG12-AS1	Reduced
cg22298224	-0.2152559	0.0388684	-5.5380682	3.06E-08	0.00083388	SART1	Reduced
cg11891887	-0.4195671	0.07600953	-5.5199273	3.39E-08	0.00088619	LINC00862	Reduced
cg08996521	-0.2957907	0.05390726	-5.4870291	4.09E-08	0.00102536	CISH	Reduced
cg02532700	-0.4236976	0.07747407	-5.4688962	4.53E-08	0.00109228	NCF4	Reduced
cg23026023	-0.2473478	0.04540939	-5.4470637	5.12E-08	0.00118941	PTPRJ	Reduced
cg07102252	-0.235349	0.04338144	-5.4251094	5.79E-08	0.00123784		Reduced
cg16531578	-0.4085217	0.07543851	-5.4152936	6.12E-08	0.00123784	SELPLG	Reduced

cg19748455	-0.2604183	0.04804406	-5.4204056	5.95E-08	0.00123784	LOC100996291	Reduced
cg21089993	-0.4014792	0.07402426	-5.4236175	5.84E-08	0.00123784	TRERF1	Reduced
cg10068788	-0.1957573	0.03621815	-5.4049494	6.48E-08	0.00127046	ASAP1	Reduced
cg07362849	-0.3743692	0.06938052	-5.395884	6.82E-08	0.00129582	BISPR;BST2;MVB12A	Reduced
cg06429483	-0.1800931	0.03361162	-5.3580613	8.41E-08	0.00155159		Reduced
cg12803068	0.76838829	0.14393859	5.33830633	9.38E-08	0.00168104	MYO1G	Increased
cg09298159	-0.3716936	0.06972609	-5.3307676	9.78E-08	0.00170367	FBXO5	Reduced
cg00841849	-0.2771429	0.0522046	-5.3087826	1.10E-07	0.00187054		Reduced
cg11148876	-0.2426246	0.0457459	-5.3037443	1.13E-07	0.00187233	RPS6KA5	Reduced
cg09731767	-0.2696584	0.05111781	-5.2752341	1.33E-07	0.00213202	RNH1	Reduced
cg00139947	0.51859301	0.09849993	5.26490715	1.40E-07	0.00219903	TOP1MT	Increased
cg02679745	-0.1955765	0.03720962	-5.2560743	1.47E-07	0.00225098	C9orf139;FUT7	Reduced
cg05009104	0.43446161	0.08288413	5.24179475	1.59E-07	0.00237446	MYO1G	Increased
cg07957619	-0.1910215	0.03650599	-5.2326057	1.67E-07	0.00243759	GTPBP2	Reduced
cg11931951	-0.1359869	0.02608139	-5.2139437	1.85E-07	0.00257635	FBLN5	Reduced
cg25132241	0.15738769	0.03017158	5.21642126	1.82E-07	0.00257635	PRAM1	Increased
cg05903736	-0.2675242	0.05152797	-5.1918259	2.08E-07	0.0027786	SHMT2	Reduced
cg08228578	-0.1344958	0.025887	-5.1954954	2.04E-07	0.0027786	HDAC4	Reduced
cg01971407	-0.1996529	0.03854964	-5.1791118	2.23E-07	0.00291282	IFITM1	Reduced
cg02695372	-0.2436274	0.04721873	-5.1595506	2.48E-07	0.00312664		Reduced
cg20962215	-0.1220912	0.02366916	-5.1582401	2.49E-07	0.00312664		Reduced
cg03221779	0.22758511	0.04434564	5.13207375	2.87E-07	0.00326754		Increased
cg11927233	-0.3739561	0.07282666	-5.1348792	2.82E-07	0.00326754	CD68	Reduced
cg15095917	-0.2715392	0.05284276	-5.1386265	2.77E-07	0.00326754	NPM1	Reduced
cg21393163	-0.4227694	0.08221494	-5.142246	2.71E-07	0.00326754		Reduced
cg26164488	-0.2756041	0.05367438	-5.1347425	2.83E-07	0.00326754	SETD3	Reduced
cg02448796	0.28437584	0.05553654	5.12051747	3.05E-07	0.00329457	HMGA2	Increased
cg07839457	-0.5659969	0.11046753	-5.1236497	3.00E-07	0.00329457	NLRC5	Reduced
cg22551157	-0.224168	0.04373775	-5.1252759	2.97E-07	0.00329457	KCNAB2	Reduced
cg00310412	-0.1656924	0.03240651	-5.1129364	3.17E-07	0.0033715	SEMA7A	Reduced
cg25780496	-0.2601229	0.05095229	-5.1052255	3.30E-07	0.00345339	LINS	Reduced
cg01657995	-0.3748057	0.07356212	-5.0950909	3.49E-07	0.00353614	C6orf48;SNORD52	Reduced
cg04804748	-0.2159343	0.04239154	-5.0938073	3.51E-07	0.00353614	LYSMD2	Reduced

cg05871405	0.24726653	0.04861545	5.08617148	3.65E-07	0.00353614	USP10	Increased
cg10213935	-0.3792702	0.0746239	-5.082423	3.73E-07	0.00353614	SH3PXD2A	Reduced
cg13289509	-0.2077117	0.04084197	-5.085742	3.66E-07	0.00353614	MOBKL2A	Reduced
cg16743273	0.32992434	0.0648952	5.08395572	3.70E-07	0.00353614	TNFSF10	Increased
cg21766592	-0.3874496	0.07627227	-5.0798227	3.78E-07	0.00353614	SLC1A5	Reduced
cg24889777	0.21447676	0.04241935	5.05610685	4.28E-07	0.00394632	LOC100505942	Increased
cg09945032	-0.2077567	0.04114414	-5.0494847	4.43E-07	0.00402638		Reduced
cg07687574	-0.2047178	0.04056952	-5.0460979	4.51E-07	0.00403982	PIK3R5	Reduced
cg22995183	0.18928651	0.037624	5.03100447	4.88E-07	0.00424982		Increased
cg24797066	-0.1616532	0.03212468	-5.0320562	4.85E-07	0.00424982	MKL2	Reduced
cg10584393	-0.1961545	0.03903559	-5.0250168	5.03E-07	0.00428879		Reduced
cg27314686	-0.1952049	0.0388545	-5.0239975	5.06E-07	0.00428879	PLAU;C10orf55	Reduced
cg13086983	-0.1975765	0.03944971	-5.0083134	5.49E-07	0.00449298	ECE1	Reduced
cg14391737	-0.3174859	0.06340314	-5.0074159	5.52E-07	0.00449298	CSF2RB	Reduced
cg26422794	-0.1985151	0.03964316	-5.0075493	5.51E-07	0.00449298	PRSS23	Reduced
cg09975715	0.29056037	0.05809046	5.00186012	5.68E-07	0.00456516	MYH10	Increased
cg11905586	-0.2676884	0.05360016	-4.9941716	5.91E-07	0.00469067	TRERF1	Reduced

Note: Top differentially methylated positions (DMPs) sorted by significance. Beta difference, standard error, test statistic, unadjusted *p*-value, FDR-adjusted *p*-value, annotated gene(s), and direction of methylation change are shown. The red dashed line marks the Bonferroni-corrected genome-wide significance threshold ($p < 9 \times 10^{-8}$).

Table 3. Frequency of exposures to potentially traumatic events during life.

Event	CUD	Controls	Chi²	p-value	Effect size (φ)
LEC-5: Death Risk	297/541 (54.9%)	132/663 (19.9%)	157.49	<0.0001	0.36
LEC-5: Other Violent Physical Harm	384/562 (68.3%)	97/663 (14.6%)	365.51	<0.0001	0.55
LEC-5: Raped	155/534 (29.0%)	18/663 (2.7%)	163.49	<0.0001	0.37
LEC-5: Witness of Extreme Violence	365/544 (67.1%)	88/663 (13.3%)	366.92	<0.0001	0.55
CTQ: Severe Emotional Abuse	141/537 (26.3%)	44/660 (6.7%)	85.47	<0.0001	0.27
CTQ: Severe Physical Abuse	146/538 (27.1%)	37/659 (5.6%)	104.29	<0.0001	0.30
CTQ: Severe Sexual Abuse	57/534 (10.7%)	19/659 (2.9%)	28.73	<0.0001	0.16
Exposure to a PTE	502/562 (89.3%)	247/663 (37.3%)	344.91	<0.0001	0.53

Note. CUD, Cocaine Use Disorder. PTE, potentially traumatic experiences. LEC-5, Life Events Checklist for DSM-5. CTQ, Childhood Trama Questionnaire.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFileHypo2.1.xlsx
- SupplementaryFileHyper2.2.xlsx
- SupplementaryFile4.xlsx
- SupplementaryFile7.xlsx
- SupplementaryFile6.docx
- SupplementaryFile8sensitivitsmkeffect.xlsx
- SupplemantalTable1.docx
- SupplementaryFile3PCA.xlsx
- SupplementaryFile5.xlsx
- SupplementaryFile1.xlsx