Aus dem Zentralinstitut für seelische Gesundheit
der Medizinischen Fakultät Mannheim
(Direktor: Prof. Dr. med. Andreas Meyer-Lindenberg)

Multi-omics characterization of cocaine use disorder in postmortem human brain

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc.hum.)
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Eric Marcel Zillich, geb. Poisel

aus
Crailsheim
2025

Dekan: Herr Prof. Dr. med. Sergij Goerdt
Referent(in): Frau Priv.-Doz. Dr. Stephanie Witt

TABLE OF CONTENTS

Page

ABBREVIATIONS

AA	African American
ACC	Anterior cingulate cortex
AMG	Amygdala
AS	Alternative splicing
AUD	Alcohol use disorder
BA46	Brodmann Area 46
BA9	Brodmann Area 9
CN	Caudate nucleus
CpG	Cytosine-phosphate-guanine
CUD	Cocaine use disorder
DAT	Dopamine transporter
DBCBB	Douglas Bell Canada Brain Bank
DEG	Differentially expressed gene
DEP	Differentially expressed protein
dlPFC	Dorsolateral prefrontal cortex
DMR	Differentially methylated region
DNAm	DNA methylation
DSM-5	Diagnostic and Statistical Manual of Mental Disorders 5
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders IV
EA	European American
EAA	Epigenetic age acceleration
EWAS	Epigenome-wide association study
FDR	False discovery rate
GO	Gene Ontology
GSEA	Gene set enrichment analysis
GWAS	Genome-wide association study
ICD-11	International Statistical Classification of Diseases 11
IEG	Immediate early gene
KEGG	Kyoto encyclopedia of genes and genomes
MDD	Major depressive disorder
MOFA	Multi omics factor analysis
MSN	Medium spiny neuron
NAc	Nucleus accumbens
NET	Norepinephrine transporter
OFC	Orbitofrontal cortex
OUD	Opioid use disorder
PCA	Principal component analysis
PFC	Prefrontal cortex
PMI	Postmortem interval
PPI	Protein-protein interaction

PUT	Putamen
RIN	RNA integrity number
RRBS	Reduced-representation bisulfite sequencing
RRHO	Rank-rank hypergeometric overlap
SERT	Serotonin transporter
SNP	Single nucleotide polymorphism
snRNA-seq	Single-nuclei RNA sequencing
SUD	Substance use disorder
TMT	Tandem mass tag
UMAP	Uniform manifold approximation and projection
vlPFC	Ventrolateral prefrontal cortex
vmPFC	Ventromedial prefrontal cortex
VS	Ventral striatum
VTA	Ventral tegmental area
WGCNA	Weighted gene co-expression network analysis

PREAMBLE

This is a cumulative dissertation based on three studies that I first-authored. The study, "*DNA methylation in cocaine use disorder-An epigenome-wide approach in the human prefrontal cortex*", published in *Frontiers in Psychiatry*, represents chapter 2 of this thesis. The second study, published in *Translational Psychiatry* with the title "*Multi-omics profiling of DNA methylation and gene expression alterations in human cocaine use disorder*" is provided as chapter 3. The third study, "*A multi-omics and cell type-specific characterization of the ventral striatum in human cocaine use disorder*", published in *Cell Reports* represents chapter 4. I independently conducted literature research, drafted and designed study and analysis plans, and performed all statistical analyses while supervised by the co-authors. Throughout the three studies, I have interpreted the results, drafted the manuscript text, and conducted the revision process according to suggestions from co-authors and from peer reviewers.

Poisel, E., Zillich, L., Streit, F., Frank, J., Friske, M. M., Foo, J. C., Mechawar, N., Turecki, G., Hansson, A. C., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2023). DNA methylation in cocaine use disorder–An epigenome-wide approach in the human prefrontal cortex [Original Research]. *Frontiers in Psychiatry*, *14*. https://doi.org/10.3389/fpsyt.2023.1075250

Zillich, E., Belschner, H., Avetyan, D., Andrade-Brito, D., Martínez-Magaña, J. J., Frank, J., Mechawar, N., Turecki, G., Cabana-Domínguez, J., Fernàndez-Castillo, N., Cormand, B., Montalvo-Ortiz, J. L., Nöthen, M. M., Hansson, A. C., Rietschel, M., Spanagel, R., Witt, S. H., & Zillich, L. (2024). Multi-omics profiling of DNA methylation and gene expression alterations in human cocaine use disorder. *Translational Psychiatry*, *14*(1), 428. https://doi.org/10.1038/s41398-024-03139-9

Zillich, E., Artioli, A., Rossetti, A. C., Avetyan, D., Belschner, H., Frank, J., Stein, F., Schwarz, J. J., Mechawar, N., Turecki, G., Nöthen, M. M., Hansson, A. C., Witt, C. C., Rietschel, M., Koch, P., Spanagel, R., Zillich, L., & Witt, S. H. (2025). A multi-omics and cell type-specific characterization of the ventral striatum in human cocaine use disorder. *Cell Reports*, *44*(2). https://doi.org/10.1016/j.celrep.2025.115332

Darstellung der Eigenleistung der Doktorandin/des Doktoranden

bei kumulativen Dissertationen

Name der Doktorandin/des Doktoranden Eric Zillich

Titel der Dissertation
Multi-omics characterization of cocaine use disorder in postmortem human brain

Betreut durch PD Dr. Stephanie Witt

☐ Ich möchte eine kumulative Dissertation einreichen und bitte den Promotionsausschuss zu prüfen, ob die vorgeschlagenen Publikationen quantitativ und qualitativ ausreichen, um die Anforderungen an eine kumulative Dissertation zu erfüllen.

☒ Der Promotionsausschuss hat zuvor geprüft, ob meine Publikationen für eine kumulative Dissertation geeignet sind, und dies ist eine abschließende Übersicht über die in meiner kumulativen Dissertation enthaltenen Publikationen.

1. Liste der peer-reviewed Publikationen, die in die kumulative Dissertation aufgenommen werden. Geben Sie für jede Publikation eine vollständige Liste der Autoren, den Titel, die Zeitschrift, den Impact Factor der Zeitschrift an und ob das Manuskript zur Veröffentlichung angenommen wurde, sich nach der Begutachtung in Überarbeitung befindet oder eingereicht wurde und zur Begutachtung ansteht. Geteilte Erstautorenschaften sollten deutlich angegeben werden. Bitte geben Sie auch an, ob es sich bei der Publikation um einen Original-Forschungsbericht, einen Review oder eine andere Art von Artikel handelt.

Publikation 1 Poisel, E., Zillich, L., Streit, F., Frank, J., Friske, M. M., Foo, J. C., Mechawar, N., Turecki, G., Hansson, A. C., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2023). DNA methylation in cocaine use disorder–An epigenome-wide approach in the human prefrontal cortex [Original Research]. Frontiers in Psychiatry, 14. https://doi.org/10.3389/fpsyt.2023.1075250, IF 3.2, published

Publikation 2 Zillich, E., Belschner, H., Avetyan, D., Andrade-Brito, D., Martínez-Magaña, J. J., Frank, J., Mechawar, N., Turecki, G., Cabana-Domínguez, J., Fernàndez-Castillo, N., Cormand, B., Montalvo-Ortiz, J. L., Nöthen, M. M., Hansson, A. C., Rietschel, M., Spanagel, R., Witt, S. H., & Zillich, L. (2024). Multi-omics profiling of DNA methylation and gene expression alterations in human cocaine use disorder. Translational Psychiatry, 14(1), 428. https://doi.org/10.1038/s41398-024-03139-9, Original Research, IF 5.8, published

Publikation 3 Zillich, E., Artioli, A., Rossetti, A. C., Avetyan, D., Belschner, H., Frank, J., Stein, F., Schwarz, J. J., Mechawar, N., Turecki, G., Nöthen, M. M., Hansson, A. C., Witt, C. C., Rietschel, M., Koch, P., Spanagel, R., Zillich, L., & Witt, S. H. (2025). A multi-omics and cell type-specific characterization of the ventral striatum in human cocaine use disorder. Cell Reports, 44(2). https://doi.org/10.1016/j.celrep.2025.115332, Original Research, IF 7.5, published

2. Zusammenfassung des Beitrags der Doktorandin/des Doktoranden zu der in jedem Manuskript berichteten Arbeit

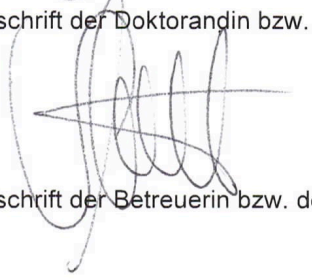
Arbeitsschritte	Publikation 1	Publikation 2	Publikation 3
Konzeption (%)	75	90	90
Literaturrecherche (%)	95	95	95
Ethikantrag (%)	30	30	30
Tierversuchsantrag (%)	-	-	-
Datenerhebung (%)	-	-	-
Datenauswertung (%)	75	95	100
Ergebnisinterpretation (%)	90	90	90
Verfassen des Manuskripttextes (%)	90	90	90
Revision (%)	90	90	90
Geben Sie an, welche Abbildungen/ Tabellen aus Ihrer Doktorarbeit entstanden sind.	alle	alle	alle
Geben Sie im Einzelnen an, welche Daten/Zahlen/Tabellen auf Forschungsergebnissen von anderen beruhen.	-	Replikationsdatensätze BA9+BA46	snRNA-seq Datensatz Ratten

3. Die Mindestanzahl der Publikationen, die für eine publikationsbasierte kumulative Dissertation erforderlich sind, ist in den "Ausführungsbestimmungen zu publikationsbasierten Dissertationen" festgelegt. Im Falle einer gemeinsamen Erstautorenschaft oder einer Letztautorenschaft begründen Sie bitte unten, warum die Veröffentlichung einer einzelnen Erstautorenschaft gleichgestellt werden soll.

-

4. Ich bestätige hiermit, dass dies eine wahrheitsgetreue Darstellung des Beitrags der Doktorandin/des Doktoranden zu den aufgeführten Publikationen ist.

Unterschrift der Doktorandin bzw. des Doktoranden



Unterschrift der Betreuerin bzw. des Betreuers

1. INTRODUCTION

1.1 Cocaine use disorder

1.1.1 Cocaine

Cocaine, or benzoylecgonine, is a tropane alkaloid naturally occurring in the *Erythroxylon coca* plant that is indigenous to Southern American countries (Gold, 1993). In countries where this plant is native, the consumption of coca leaves from the *Erythroxylon* plant has a centuries-long socio-cultural history. In contrast, the use of pure cocaine, depicting the primary psychoactive compound of coca leaves, represents a more recent phenomenon (Biondich & Joslin, 2016). Cocaine mainly refers to cocaine hydrochloride powder that is consumed intranasally, but also injections of aqueous solutions of cocaine hydrochloride are reported (Saunders et al., 2016). Another form of cocaine is crack cocaine, the free cocaine base, which is derived from a chemical reaction of cocaine hydrochloride with sodium bicarbonate (Smart, 1991). Crack cocaine is usually smoked, and due to its application route, it leads to a particularly rapid onset of cocaine effects. Cocaine acts as a potent psychostimulant, with euphoria, hyperactivity, and increased risk for aggressive behavior as the typical acute psychotropic effects of cocaine that have, compared to other psychostimulants, a relatively short duration of around 20-30 minutes (Saunders et al., 2016; Schwartz et al., 2022). Further, cocaine has strong vasoconstrictive but also local anesthetic properties (Lange et al., 1989; López-Valverde et al., 2014). At the molecular level, cocaine blocks neurotransmitter reuptake transporters in the synaptic cleft, such as DAT, SERT, and NET, thereby acting as a reuptake inhibitor in the brain (Hall et al., 2009; Ritz et al., 1987). The increased levels of neurotransmitters remaining in the synaptic cleft are assumed to cause the rewarding effects of cocaine by interfering with synaptic signaling, particularly of dopaminergic neurocircuits (Hall et al., 2009).

1.1.2 Symptoms and diagnostic criteria

Individuals who repeatedly use cocaine can develop cocaine addiction, characterized by a strong compulsion to seek and take cocaine, difficulties in controlling cocaine intake, and negative emotional symptoms when cocaine is not used (Koob & Volkow, 2016). Cocaine addiction presents as a chronically relapsing disorder and is described in different diagnostic systems. The diagnosis of cocaine use disorder (CUD) refers to the substance use disorder nomenclature of the Diagnostic and Statistical Manual of Mental Disorders (5th edition, DSM-5) in which CUD is described as a subcategory of stimulant use disorders. The DSM-5 reports 11 diagnostic criteria for CUD, of which at least two have to be fulfilled within a 12-month period in order to diagnose a CUD (American Psychiatric Association, 2013). These diagnostic criteria include, for instance, cocaine craving,

the inability to reduce or limit cocaine consumption, continued cocaine use despite negative consequences, and the development of tolerance and withdrawal. In the DSM-5, three severity grades of CUD are described. Mild CUD corresponds to the fulfillment of 2-3 symptoms, moderate to the presence of 4-5, and severe to the presence of 6 or more symptoms.

In contrast, the most recent 11th version of the International Statistical Classification of Diseases and Related Health Problems (ICD-11) differentiates between harmful use of cocaine and cocaine dependence (World Health Organization, 2019). Harmful use describes cocaine use patterns that have a negative influence on the individual's but also on other's mental or physical health. This includes the consequences of intoxication-related behavior, toxic effects on organs and tissues, and harmful administration routes of cocaine. Cocaine dependence, according to ICD-11, is characterized by three major symptom categories: 1) impaired control over cocaine intake, often associated with a strong urge to use cocaine; 2) physiological/adaptive symptoms such as the development of tolerance and withdrawal; and 3) the prioritization of cocaine use over other activities in life and sustained cocaine use despite negative consequences. To diagnose cocaine dependence, two or more of these three criteria have to be fulfilled over a period of 12 months or over a period of 1 month if cocaine is consumed daily or almost daily. To achieve a consistent terminology throughout this thesis, the term cocaine use disorder according to DSM-5 will be used in the following.

1.1.3 Epidemiology

The development of CUD follows a dynamic process characterized by a gradual increase in symptom severity and loss of control over cocaine intake (Piazza & Deroche-Gamonet, 2013). Dependent on drug expectancies and its availability, i.e., exposure opportunity, the first use of cocaine is often of an experimental nature (Figure 1.1). After the initial contact with cocaine, a recreational cocaine use pattern may be developed if the psychotropic and rewarding effects of cocaine are perceived as pleasant. Here, periods of abstinence and control over cocaine intake can be maintained while also cessation of drug use is observed at this stage (Bogdan et al., 2023). However, a substantial fraction of ~20% of regular cocaine users is assumed to transition to CUD (Lopez-Quintero et al., 2011). While cessation still exists as a theoretical possibility in CUD, it is, in fact, extremely difficult to limit or even stop cocaine consumption due to the vicious symptom cycle, including intoxication, withdrawal, and craving (Figure 1.1).

Risk factors for the development of CUD include genetic and environmental components. Heritability estimates for CUD range between 65% and 79% (Kendler

et al., 2000; Kendler et al., 2007), indicating that the genetic component is substantial for CUD compared to other SUDs, such as alcohol use disorder (AUD), with an estimated heritability of around 50% (Verhulst et al., 2015). However, there are also clear environmental factors, such as drug availability, household drug use, and peer drug use, that are associated with an increased risk of developing CUD (Pierce et al., 2018). Also, sex differences were reported in CUD, with women showing stronger vulnerability to different aspects of CUD (Knouse & Briand, 2021). For instance, women reported stronger cocaine craving and used cocaine more frequently than previously intended compared to men. Such sex-specific effects were also observed in animal models of CUD, which identified ovarian hormones as important biological drivers of sex differences in CUD (Evans & Foltin, 2010; Lynch, 2008). Further, ethnic background was found to influence the disease course of CUD. For instance, in African American populations, a higher consumption rate of crack cocaine was found, and also significantly fewer treatment contacts were observed during a 6-month follow-up period (Bernstein et al., 2006). These findings suggest a complex interplay between genetic and environmental factors that modulate the risk of developing CUD.

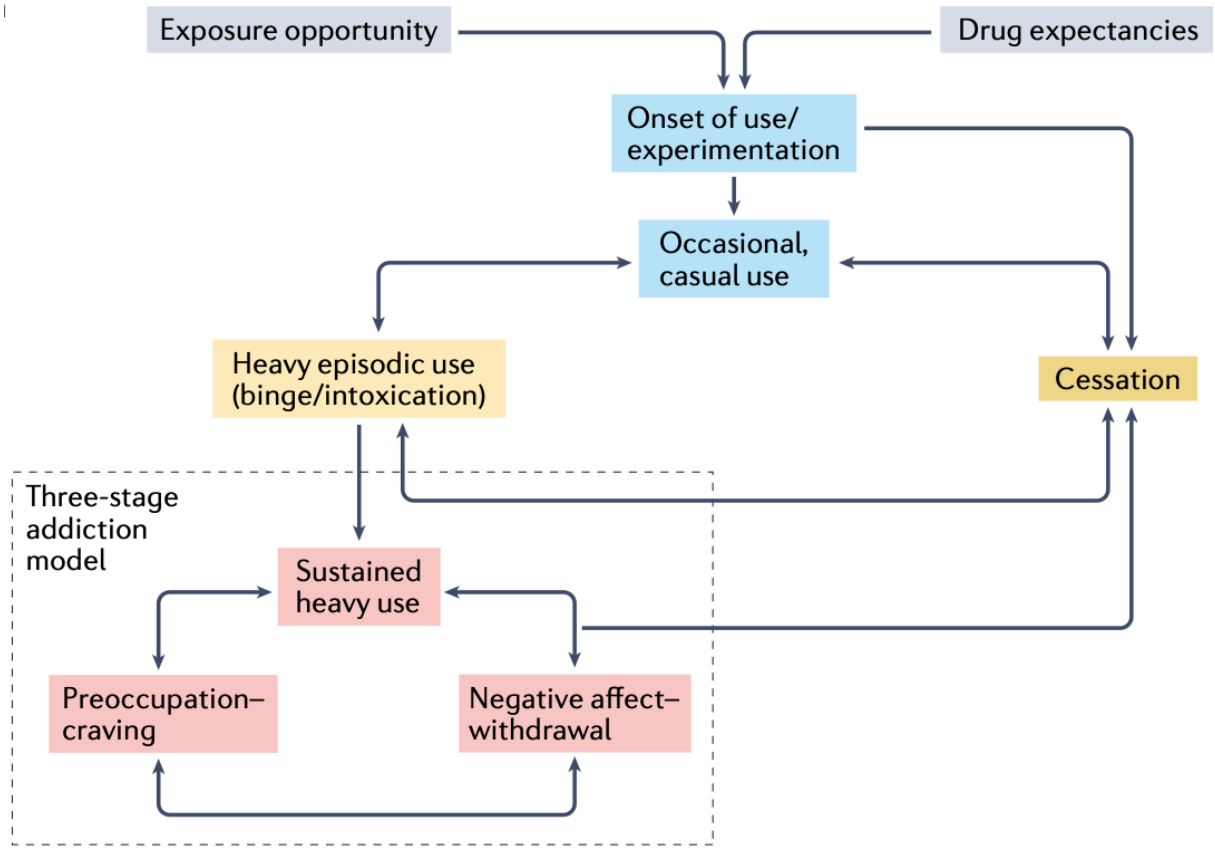


Figure 1.1 - A multi-step process is assumed to underlie the development of addiction. It is initiated by the first contact with the drug of abuse, often consumed experimentally. After occasional and repeated drug use, a transition to heavy episodic use and addiction can follow based on individual vulnerability influenced by genetic and environmental factors. While cessation of drug use depicts a potential trajectory throughout all stages, it is particularly challenging as soon as the transition to addiction occurs. Figure from Bogdan et al. (2023), reprinted with permission from Springer Nature.

The prevalence of cocaine consumption and CUD displays profound geographical differences, which are strongly associated with differences in cocaine markets and, therefore, cocaine availability. While the majority of cocaine production is located in Southern American countries where coca bushes are cultivated, the strongest cocaine markets are observed in North American and Western/Central European countries (UNODC, 2023a). Small but detectable markets were reported for Asian and African countries. During the last decade, a sharp increase in cocaine production and its market has been observed, which led to a global increase in cocaine users that even exceeds the population growth rate. Globally, a total of 21.5 million people used cocaine in 2020, and estimates range from 25 to 55 million additional cocaine users in the coming years if the growing markets also reach the previously underrepresented populations (UNODC, 2023a, 2023b). Prevalence estimates for past-year cocaine use in North American and Western/Central European countries range between 1% and 2.5% (UNODC, 2023a), while for cocaine use disorder, for example, in the United States, the prevalence is estimated to be around 0.37-0.71% (Mustaquim et al., 2021). Cocaine consumption puts individuals at considerable risk for developing CUD, as not only a substantial fraction of around 20% of regular cocaine users transition to CUD during their lifetime, but also transition rates of ~15% from use to use disorder within 10 years are found for cocaine, which exceeds transition rates observed for alcohol (~12%) and cannabis (~8%) (Wagner & Anthony, 2002). Further, even from the first use of cocaine, transition rates to use disorder during a 1-year period were estimated to be around 7,1% for cocaine, which far exceed those of alcohol, cannabis, and nicotine (~2%) (Lopez-Quintero et al., 2011). Cocaine thus depicts a particularly hazardous drug, posing individuals at significant risk for developing a substance use disorder.

The repeated consumption of cocaine in individuals with CUD is strongly related to negative health outcomes due to the direct and indirect effects of cocaine exposure. Lethal cocaine intoxications depict the probably most detrimental negative health outcome, with more than 27,000 individuals in the US alone dying with or due to a cocaine overdose in 2022, where cocaine was often consumed together with opioids (NIDA, 2024). Further, the risk for ischemic heart attack (Mittleman et al., 1999) and stroke (Cheng et al., 2016), other common causes of death, is strongly increased in individuals with CUD. Next to negative physical consequences, CUD is also associated with negative psychosocial outcomes such as social isolation, loss of employment, and an increased risk of poverty (Daley, 2013). There is also an increased risk for the development of comorbid psychiatric disorders in CUD, including psychotic (Sabe et al., 2021) and mood disorders (Ford et al., 2009) that further exacerbate the negative effects on health. It has to be noted that individuals with CUD often present with polysubstance use, i.e., the consumption of substances such as nicotine, cannabis, alcohol, and opioids in addition to cocaine (Stiltner et al., 2023). This leads to an increased likelihood for

the co-occurrence of multiple substance use disorders, further increasing the risk for drug-related incidents and potentiating the long-term negative consequences of (poly-) substance use (Liu et al., 2018).

1.1.4 Treatment

In contrast to other SUDs, such as AUD, there is currently no FDA-approved medication for the treatment of CUD (Schwartz et al., 2022). This means that psychosocial therapy currently depicts the gold standard for treatment-seeking patients (Saunders et al., 2016). Different psychotherapy approaches, such as intensive outpatient therapy (IOT), cognitive behavioral therapy (CBT), and voucher-based reinforcement therapy (VBRT), are available (Kampman, 2019). IOT consists of individual and group counseling sessions and was shown to be effective in reducing the days of cocaine use by around 50% (Coviello et al., 2001). While VBRT, following the idea of contingency management to promote the intended behavior, i.e., abstinence (Higgins et al., 2000) is particularly effective in promoting abstinence from cocaine use, CBT was shown to be more protective against relapse episodes (Maude-Griffin et al., 1998). High costs for psychosocial therapies, substantial drop-out rates, and relapse in a majority of ~75% of patients (Paliwal et al., 2008) depict significant challenges in the treatment of CUD.

While there is currently no approved pharmacotherapy against CUD, several drug candidates have been evaluated in clinical studies. Agonist replacement therapies similar to methadone in opioid use disorder and varenicline in tobacco use disorder have been tested in CUD, for instance, with long-acting amphetamine, resulting in positive results on the reduction of cocaine intake (Mooney et al., 2009; Nuijten et al., 2016). Modulation of GABAergic and glutamatergic signaling by topiramate was shown to be effective in increasing abstinence days in individuals with CUD in clinical trials (Baldaçara et al., 2016; Johnson et al., 2013; Kampman et al., 2013), however, only during early follow-up. In addition, negative results on topiramate were published (Nuijten et al., 2014). In a recent phase II clinical trial on mavoglurant, an mGluR5 glutamate receptor antagonist, participants in the mavoglurant group reported significantly fewer cocaine use days ($p=0.021$) compared to placebo (Gomez-Mancilla et al., 2025) suggesting the glutamatergic system as a promising drug target in CUD. Thus, while there are promising results on potential pharmacotherapies for the treatment of CUD, further research on molecular alterations in CUD is required to develop targeted and mechanism-based interventions.

1.2 Neuroanatomical insights into cocaine use disorder

1.2.1 The neurocircuitry of addiction

The brain is assumed to be the primary organ involved in the development and maintenance of SUDs, such as CUD (Volkow et al., 2016). Due to advances in human neuroimaging and animal models of addiction, specific neurocircuits have been identified that are associated with the three key stages of the addiction cycle: the binge/intoxication phase, the negative affect/withdrawal phase, and the preoccupation/anticipation stage (Koob & Volkow, 2010; Volkow et al., 2003). Each stage of the addiction cycle has been assigned to a set of connected brain regions that are particularly relevant for symptom expression, thereby forming a neurocircuitry of addiction, as shown in Figure 1.2 (Koob & Volkow, 2016).

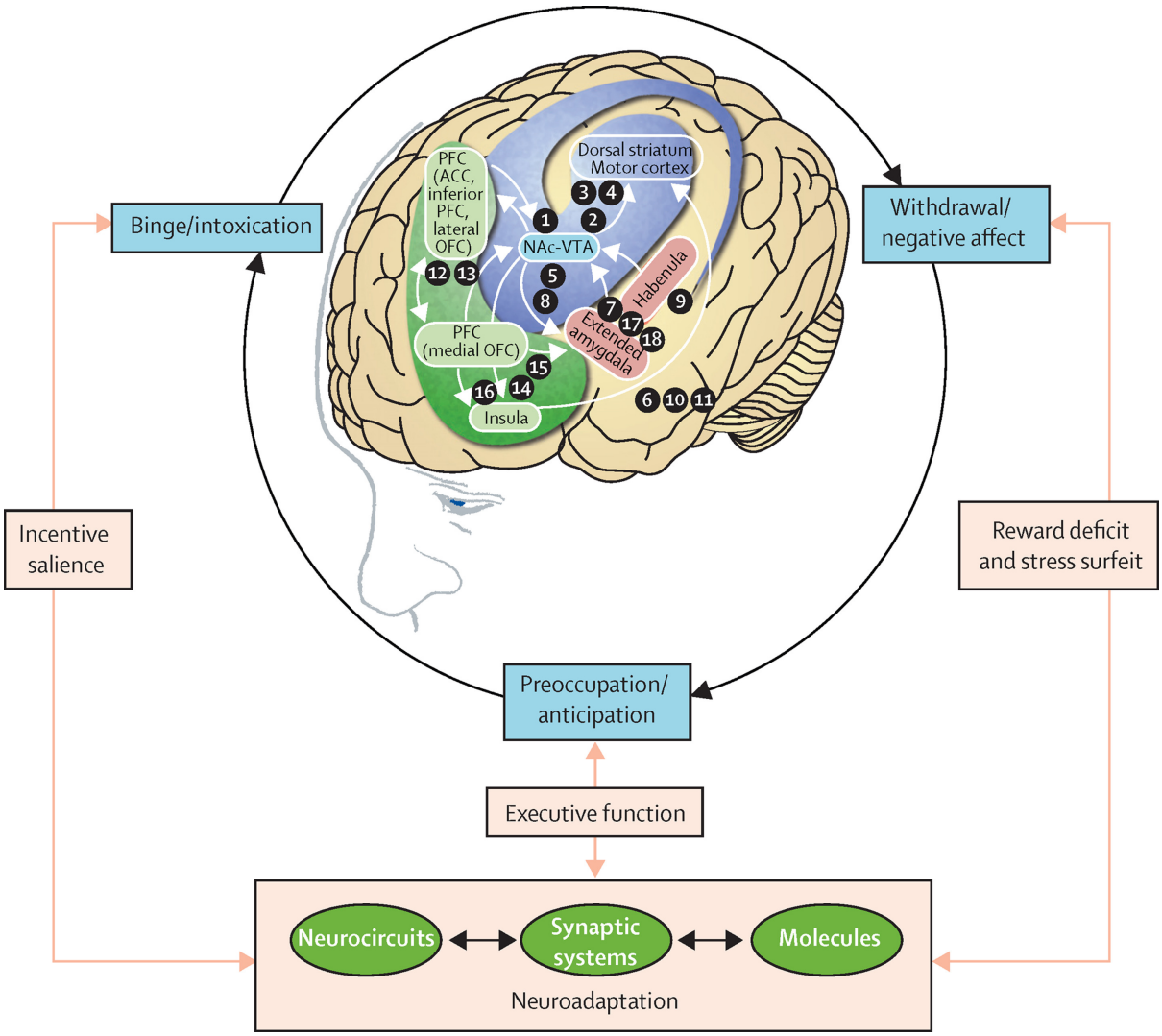


Figure 1.2 - The neurocircuitry of addiction. Based on cumulative evidence from human neuroimaging studies and animal models of addiction, a conceptual framework has been developed that describes a set of neurocircuits in the brain that are associated with the three different stages of the addiction cycle. Molecular and synaptic system adaptations are assumed to provide a mechanistic basis for neurocircuit alterations in addiction. According to this conceptual framework,

studying neuroadaptations at multiple levels in the brain depicts a key research area to better understand the neurobiological processes underlying addiction. Figure from Koob and Volkow (2016), reprinted with permission from Elsevier.

The binge/intoxication stage primarily involves striatal brain regions (Everitt & Robbins, 2013). More specifically, ventral tegmental area (VTA) dopaminergic neurons project to the nucleus accumbens (NAc), depicting the part of the ventral striatum whose activation is strongly associated with reward perception (Pignatelli & Bonci, 2015). The dorsal striatum, consisting of the putamen (PUT) and caudate nucleus (CN), is associated with habit formation and compulsive drug seeking, particularly observed at later stages of addiction (Belin & Everitt, 2008). The second stage, negative affect/withdrawal, is characterized by dysphoria, irritability, and loss of interest in non-drug-related rewards. Here, neuroadaptations in response to chronic activation of reward-related brain regions lead to altered neurochemical states of the brain as soon as the drug is not consumed anymore, inducing a negative emotional state (Koob & Volkow, 2016). The negative affect/withdrawal stage is associated with elevated levels of corticotropin-releasing factor and corticosterone in the amygdala (AMG) (Koob et al., 2014) and reduced dopamine levels in the NAc (Weiss et al., 1992). These negative emotional states are potent inducers of negative reinforcement, as drug use reactivates the reward system, however, at the cost of producing further neuroadaptations enforcing anti-reward circuits (Koob & Volkow, 2016). Lastly, the preoccupation/anticipation stage is related to craving and relapse in SUDs (Koob, 2008). Glutamatergic projections from frontal cortical regions to the VTA and striatal regions are assumed to ensure executive control, thereby preventing excessive drug-taking in the presence of incentive-salient drug-related cues (Goldstein & Volkow, 2011; Hester & Garavan, 2004). This system is altered in SUDs, thereby enabling excessive drug intake. A prefrontal, glutamatergic system that is assumed to underlie craving during the preoccupation/anticipation stage has been described (Frye et al., 2016; Kalivas et al., 2005). Further, it was shown that a hypoactive frontal cortex is associated with executive function deficits such as reduced inhibitory control in SUDs (Goldstein & Volkow, 2011). Molecular and cellular adaptations of the brain in response to chronic drug use depict an important mechanism that underlies the complex neurocircuit alterations in SUDs (Nestler et al., 1996). As these adaptations are a consequence of genetic, epigenetic, and transcriptional mechanisms (Nestler & Lüscher, 2019), a better understanding of the molecular mechanisms underlying SUDs is crucial to identifying novel interventions and treatment options for these disorders.

1.2.2 Striatum

In humans, the striatum consists of multiple subcortical nuclei closely related to the limbic system, motor control, and reward processing (Pauli et al., 2016). Due

to its close relationship to reward and emotional processing, the striatum has been recognized early on as an important brain region in addiction, and intensive research has been performed on the neuroanatomy and function of the striatum in SUDs (Everitt & Robbins, 2013; Yager et al., 2015). The striatum can be subdivided into a ventral striatum, consisting of the nucleus accumbens (NAc), and the olfactory tubercle, while the dorsal striatum consists of CN and PUT. As the ventral striatum receives dopaminergic input from the VTA via the mesocorticolimbic system, it is particularly important for mediating the reinforcing effects of drugs of abuse (Pignatelli & Bonci, 2015). Increased dopamine levels in the NAc, as observed during cocaine intoxication, activate dopamine receptors, particularly on medium spiny neurons (MSNs) that account for 90-95% of neurons in the striatum (Yager et al., 2015). The activity of striatal MSNs is further modulated by glutamatergic projections, for instance, from the cortex and thalamus (Yager et al., 2015). The dorsal striatum is involved in basal ganglia circuits, exerting a vital motor control function. It is also related to the development of addiction, as reinforcement learning and habit formation depict important processes related to dorsal striatal brain regions that are assumed to contribute to the formation of compulsive cocaine consumption patterns (Everitt & Robbins, 2013). During the disease course of SUDs, a shift in activity was observed from the ventral to the dorsal striatum (Koob & Volkow, 2016). This is assumed to reflect the shift from initial goal-directed to the habitual use of drugs that is characteristic of later stages of SUDs (Vollstädt-Klein et al., 2010). Thus, changes in striatal neurocircuits depict a critical feature of SUDs, suggesting the striatum as a particular area of interest for studying brain changes in CUD.

1.2.3 Prefrontal Cortex

The prefrontal cortex (PFC) depicts an important brain region in CUD due to its role in regulating cocaine intake and its close relationship to executive functions such as cognitive control (Goldstein & Volkow, 2011). CUD is characterized by excessive cocaine intake that often occurs despite the knowledge about the harmful effects of cocaine. Neuroimaging studies have repeatedly identified PFC dysfunction in SUDs, leading to theoretical frameworks such as the impaired response inhibition and salience attribution (iRISA) model (Goldstein & Volkow, 2002), which is well supported by experimental data. As part of the neocortex, the PFC is an evolutionarily relatively young brain region that contains different types of excitatory and inhibitory neurons, including pyramidal neurons and interneurons, which are arranged in a characteristic six-layer architecture. PFC subregions include ventrolateral (vlPFC), ventromedial (vmPFC), and dorsolateral (dlPFC) areas, as well as the orbitofrontal cortex (OFC) and frontomedial areas, including the anterior cingulate cortex (ACC) (Goldstein & Volkow, 2011). SUDs are broadly associated with a decreased PFC volume, which was shown to

be a consequence of chronic drug intake (Jedema et al., 2021). This volume reduction was partly reversible after cessation, suggesting neuroplastic changes in the PFC to correlate with SUDs and recovery (Connolly et al., 2013). Increased activity of the dlPFC, ACC, and OFC is a characteristic feature of cocaine intoxication (Howell et al., 2010), but also cocaine-related cues alone were shown to produce similar PFC activations along with strong cocaine craving (Garavan et al., 2000). In addition, cognitive tests showed reduced inhibitory control in individuals with CUD related to ACC and dlPFC dysfunction (Li et al., 2008). An impaired disease awareness frequently observed in individuals with SUDs was further related to PFC dysfunction, particularly of the ACC (Moeller et al., 2014). In summary, the PFC is involved in several higher-order cognitive functions dysregulated in CUD, such as inhibitory control, cue reactivity, reward-related decision-making, and the ability to recognize illness. Investigating the PFC is thus of particular importance to better understand how deregulation of this brain region is involved in the pathophysiology of CUD.

1.3 Defining the molecular architecture of cocaine use disorder

1.3.1 Genetics of cocaine use disorder

Genetic variation is an inherent driving force of evolution, which, through its effect over many generations, enables improved adaptation to environmental conditions. However, a genetic contribution to the manifestation of many diseases has been proven, suggesting individual genetic variability as an important factor that determines disease susceptibility. Twin studies and heritability analyses have shown a substantial genetic risk in the development of SUDs, including CUD (Deak & Johnson, 2021). In CUD, a heritability of 0.65-0.79 was estimated (Kendler et al., 2000; Kendler et al., 2007). Further, as is true for most psychiatric disorders, CUD is a polygenic disorder, meaning that it is not the presence of one genetic risk variant but the interaction of many risk variants that influences the manifestation of CUD (Fernàndez-Castillo et al., 2022). To investigate and characterize genetic risk variants that are associated with CUD, genome-wide association studies (GWAS) depict a valuable methodological tool.

1.3.1.1 Genome-wide association studies of cocaine use disorder

In GWAS, association analyses of common genetic variants with the phenotype of interest are performed at the genome-wide scale. For this, genotyping of single nucleotide polymorphisms (SNPs) is performed in individuals with the phenotype of interest and control individuals, followed by association analysis of SNP genotype with the phenotype. Genome-wide variant data is often derived from microarrays that capture more than 2 million SNPs distributed across the genome.

This type of study is especially useful to characterize genetic risk factors in polygenic diseases, where, in contrast to monogenic diseases, many genetic risk variants contribute to the expression of the phenotype (Uffelmann et al., 2021). Due to the strong multiple testing burden and small effect sizes of individual variants, GWAS relies on large sample sizes, ideally in the thousands or tens of thousands, to obtain sufficient statistical power to detect genome-wide significant associations. Meta-analyses of individual cohorts are essential to achieve these sample sizes. So far, two larger GWAS meta-analyses have been published, both focusing on DSM-IV cocaine dependence, the first investigating a total of N=5,697 subjects from European American (EA) and African American (AA) ancestry as the discovery set (Gelernter et al., 2014) and the second meta-analyzing four cohorts of EA ancestry, resulting in a total of N=6,378 individuals (Cabana-Domínguez et al., 2019). Both studies report only suggestive associations, as none of the top findings were associated with CUD at genome-wide significance. Results from Gelernter et al. (2014) suggested *FAM53B*, *NCOR2*, and *CDK1* as risk genes for cocaine dependence, while *HIST1H2BD* was the top finding at the gene level in the study from Cabana-Domínguez et al. (2019). A substantially larger GWAS on CUD in N=574,127 individuals is currently in preparation and available as a conference abstract (Deak et al., 2023). However, the results have not been published yet. In this large-scale meta-analysis, 27 genome-wide significant associations mapping to the *DRD2* locus were identified in the EA meta-analysis, and 3 significant loci with *HS6ST3* emerged as the top finding in the AA meta-analysis. In the cross-ancestry meta-analysis, the top association was again mapped to the *DRD2* locus, while 11 significant loci were identified. Enrichment analyses revealed enrichment within synapse function pathways and brain-expressed genes, suggesting important neurophysiological functions of CUD risk genes.

1.3.2 Epigenetics of cocaine use disorder

1.3.2.1 Epigenetic modifications

Epigenetics refers to the regulation of gene expression levels that is independent of changes in the DNA sequence itself (Gibney & Nolan, 2010). Instead, epigenetic modifications include covalent modifications of the DNA, such as methyl groups attached to DNA bases and histone modifications that influence the accessibility of the DNA for the transcriptional machinery (Allis & Jenuwein, 2016). Further, non-coding RNAs and chromatin remodeling enzymes, such as the SWI/SNF complex, are considered epigenetic regulators. One of the best-studied epigenetic modifications is DNA methylation (DNAm). Here, a methyl group is attached to DNA bases, most prominently at the C5 of cytosine, resulting in 5-methylcytosine that is often enriched at gene promoters where it has an essential role in regulating gene expression. Cytosine methylation is particularly abundant at cytosines

adjacent to a guanine base, forming a so-called CpG site. These CpG sites show an uneven distribution across the genome and form CpG islands, of which more than 28 million are known in the human genome (Lövkvist et al., 2016). While cytosine methylation at promoter sites is often considered to be transcriptionally repressive, a strong context dependence of methylation marks on gene expression was shown, suggesting DNA methylation as a complex epigenetic regulation mechanism (Ambrosi et al., 2017). Environmental factors strongly influence DNAm levels, as induction or repression of gene expression depicts an important adaptation mechanism to changes in the environment, highlighting epigenetics as an important interaction point between the genome and the environment. In humans, enzymes involved in transferring methylation groups to DNA bases include DNA methyltransferases (DNMTs), while demethylation enzymes such as TET demethylases gradually oxidize and remove the methylation marks. Studying epigenetics thus contributes to a better understanding of disease mechanisms, as it provides crucial insights into how the interaction between genetic and environmental factors shapes the development and progression of diseases.

1.3.2.2 Epigenome-wide association studies of cocaine use disorder

Similarly to GWAS, which investigate the association of genetic variants with a phenotype, in epigenome-wide association studies (EWAS), association analyses of DNAm levels with the phenotype of interest are performed. Profiling of DNAm levels is often performed using microarray techniques such as the Illumina HumanMethylationEPIC BeadChip that covers more than 850,000 CpG sites across the human genome. While, in principle, whole-genome and reduced-representation bisulfite sequencing (RRBS) methods are available that can profile substantially more CpG sites than arrays, there is still an important role for array solutions in large-scale EWAS as they allow cost-efficient and high-throughput sample processing. DNAm levels are strongly tissue-dependent and are also heavily influenced by environmental factors (Feil & Fraga, 2012). Therefore, EWAS require tissue-specific analyses and rigorous adjustment for covariates such as cell type distribution, sex, age, or dietary factors (Michels et al., 2013).

Blood is easily accessible in living human beings and interacts with almost all tissues of the human body. EWAS are thus often performed in blood with the assumption that blood epigenetics serves as a surrogate for epigenetic processes that also affect other organs. Stable epigenetic associations were identified with SUDs in blood, such as differential methylation of *AHRR* CpG sites in tobacco smoking (Joehanes et al., 2016). In CUD, an EWAS in blood was performed in a cohort of N=47 individuals (N=23 CUD, N=24 controls) revealing 186 differentially methylated positions enriched in genes that contribute to transcription and chromatin regulation (Camilo et al., 2019). While there is evidence that SUDs are

systemic disorders and are therefore not only disorders of the brain, blood DNAm levels might not sufficiently reflect disease processes in other organs such as the brain, highlighting the need to also study epigenetics directly in other tissues. Postmortem human brain tissue represents a valuable resource for studying epigenetics and other strongly tissue-specific processes, such as gene expression, in the brain. Two genome-wide epigenetic studies on DNAm signatures of CUD were performed, one in the CN (Vaillancourt, Yang, et al., 2021), and the other in the NAc (Vaillancourt, Chen, et al., 2021). In the CN, RRBS was performed in a discovery cohort of N=50 male individuals, identifying 6712 differentially methylated CpG regions containing at least 2 CpG sites and 173 differentially methylated regions (DMRs). Among genomic regions with the strongest differential methylation in CUD, the *IRXA* gene cluster was identified. Follow-up analyses revealed hypomethylation in *IRX2* exon 3 to be specific to neurons. The *IRX2* DMR was identified to harbor a CTCF binding site, which induced differential gene expression by altering 3D chromatin conformation (Vaillancourt, Yang, et al., 2021). In the study on CUD-associated differential methylation in the NAc, RRBS in the same cohort size of N=50 revealed hypermethylation of a DMR in the tyrosine hydroxylase (*TH*) gene as one of the key findings (Vaillancourt, Chen, et al., 2021). The DMR in *TH* contained a putative EGR1 binding site representing a well-described cocaine-induced IEG. Together, these studies provide important insights into DNAm differences in the brain and, by performing functional validation experiments, suggest mechanistic models of how methylation changes impact gene expression in CUD.

1.3.3 Gene expression in cocaine use disorder

1.3.3.1 Gene expression – from DNA to RNA to proteins

Gene expression refers to the transfer of genetic information into RNA that is further translated to the protein level. In the first step, DNA is transcribed into mRNA by RNA polymerases that require promoter sequences to initiate the transcription of adjacent genes. At promoter regions, extensive regulation of gene expression is observed, not only by DNAm but also by regulation of the chromatin structure, for instance, by histone modifications (Miller & Grant, 2013). Further, long-range interactions involving enhancer or silencer sequences modulate the transcriptional activity. In eukaryotes, the protein-coding exonic sequences on the transcribed pre-mRNA are interrupted by non-coding intronic sequences, which are removed by the spliceosome in a process called RNA splicing. A single pre-mRNA can be spliced into different mature mRNA sequences by alternative splicing, which leads to the formation of different protein isoforms and thus increases the variability and functionality of the protein pool (Marasco & Kornblihtt, 2023). After the formation of the mature mRNA, further regulation, for

example by microRNAs (miRNAs) comes into play, which can induce mRNA degradation based on sequence complementarity. Finally, after mRNA translation into proteins at the ribosome, post-translational modifications such as protein phosphorylation, acetylation, or ubiquitination regulate protein function and turnover, further adding complexity to the regulation of gene expression levels.

1.3.3.2 Transcriptome- and proteome-wide studies of cocaine use disorder

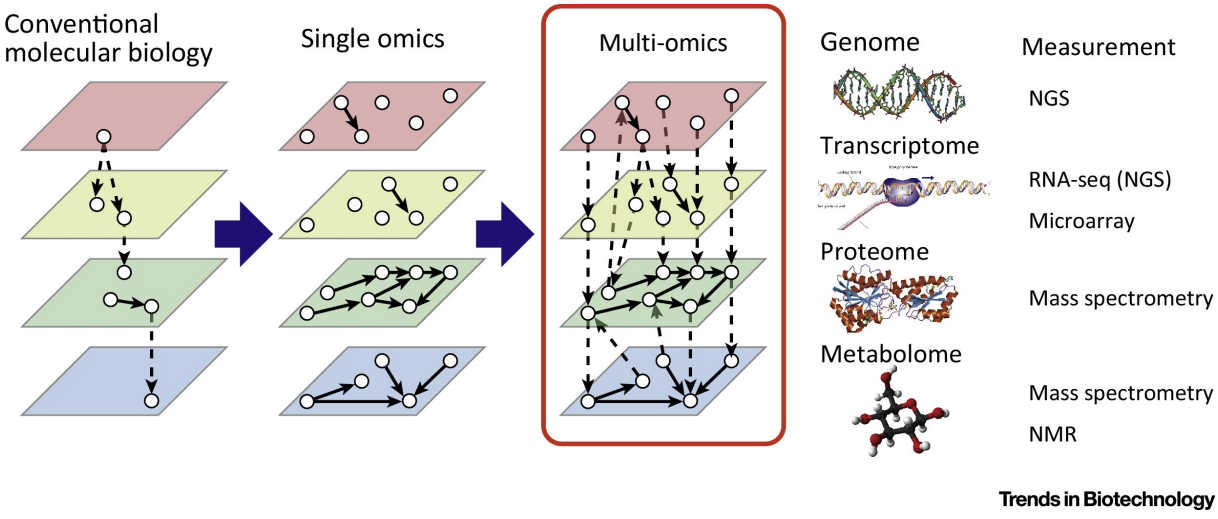
An important role of cocaine exposure in the induction of gene expression, particularly of immediate early genes (IEGs) such as CREB, JUN, and EGR1, has been described early on (Hope et al., 1992; Moratalla et al., 1993; Thiriet et al., 2000). These transcriptional changes were assumed to be induced by increased neuronal activity and represent cocaine-induced neuroadaptations at the molecular level that can result in long-term changes in neuronal function and correlate with morphological changes in dendrites and dendritic spines (Robinson et al., 2001; Volkow et al., 2019). As IEGs often act as transcription factors that can induce or repress the expression levels of their numerous target genes, a strong effect of chronic cocaine exposure on the transcriptome is expected. Several studies investigated transcriptomic changes in CUD and confirmed expression changes in a large fraction of the transcriptome. For instance, Zhou et al. (2011) reported CUD-associated transcriptional changes in postmortem hippocampus in a cohort of N=16 individuals (8 CUD, 8 controls). At 5% FDR significance, 80 differentially expressed genes (DEGs) were detected involved in mitochondrial respiration and chromatin modification. They further profiled H3K4me3, an activating chromatin mark, using ChIP-seq in the same cohort and found differential H3K4me3 distribution in promoter regions in CUD. However, no significant overlap was detected between genes with significant H3K4me3 changes and DEGs. Another study performed RNA-seq in neuronal nuclei from postmortem human brain samples of the Brodmann Area 46 subregion of the PFC in a cohort of N=36 individuals (19 CUD, 17 controls) (Ribeiro et al., 2017). The authors found N=883 DEGs at nominal significance and performed network analyses that revealed altered neuroplasticity gene networks associated with CUD. Mews et al. (2023) extended the characterization of CUD-associated transcriptional changes to striatal regions in their study of the NAc and CN in a sample of N=45 (25 CUD, 20 controls). The authors found more than 1000 DEGs at nominal significance in both brain regions and identified several CUD-associated co-expression networks, for instance, related to ion channel activity. Convergent DEG signatures were discovered between human CUD and a rodent model on chronic cocaine intake, highlighting the value of rodent models for studying the molecular underpinnings of addiction. While RNA-seq depicts an established method in SUD research, only a few proteomic studies have been performed so far. In CUD, there is currently one study that investigated proteomic changes in Brodmann Area 9 postmortem brain tissue from N=20 individuals (8 CUD, 12 controls). A total of 87 differentially

expressed proteins were found to be involved in vesicle trafficking and cell adhesion (Tondo et al., 2021). In summary, while several studies investigate gene expression differences associated with CUD in the brain, further studies are required to better understand regulatory mechanisms and the interrelationship between epigenomic, transcriptomic, and proteomic changes.

1.4 Multi-omics for an in-depth molecular characterization of cocaine use disorder

1.4.1 Definition and methods

To understand and predict the behavior of complex biological systems, a detailed and comprehensive characterization of its features is of critical importance. The suffix "ome" refers to the entirety of a specific level of features such as genes or RNAs. For instance, the entire genetic information of an organism is termed its genome, and its entire set of phenotypic characteristics is called its phenome. The better these omes are profiled using "omics" methods, the more comprehensively the system can be described. Multi-omics approaches in biological research aim to characterize and investigate multiple feature levels, such as the genome, the epigenome, the transcriptome, and the proteome, in systems such as cells, tissues, or organisms. As there is often an interrelationship between different omics levels, as with the transcriptome, which depends on the genome and the epigenome, multi-omics analyses are essential to better model complex relationships in a system where important information would not have been accessible with single-omics approaches. Single-omics analyses can only produce results that are of a correlational nature; for instance, an association between feature expression and disease status can be derived. Using multi-omics approaches, causal relationships can be inferred when information from multiple omes is related to each other (Hasin et al., 2017) (see Figure 1.3).



(Figure legend on next page)

Figure 1.3 - Comparison of conventional molecular biology approaches with omics-wide studies of molecular feature levels such as genes, transcripts, and proteins. Conventional molecular biological studies often focus on mechanistic studies investigating the role of single genes across different feature levels. In contrast, omics-wide approaches aim to capture the entirety of features of a specific level. Analyses in single omics datasets allow to the identification of within-omics relationships between features, while multi-omics, in addition, allow the identification of across-omics relationships and therefore also provide important mechanistic insights. As omics analyses are in principle not informed by previous hypotheses on disease mechanisms, as they study the entirety of features, they are hypothesis-generating by nature. Follow-up experiments can help to further understand how the findings of omics-wide analyses contribute to disease mechanisms. Figure from Yugi et al. (2016), reprinted with permission from Cell Press. The original figure was adapted as the term "Trans-omics" was replaced by "Multi-omics".

Multi-omics analyses can be performed to address different objectives in biological research. For instance, Athieniti and Spyrou (2023) reviewed the multi-omics literature and identified five distinct areas of research for which multi-omics approaches depict valuable analytical tools. These include detecting disease-associated molecular patterns, subtype identification, diagnosis and prognosis, drug-response prediction, and identifying regulatory processes. As multi-omics datasets are high-dimensional, they require sufficient computational power during data handling and analyses. Advances in bioinformatics resulted in various methods stemming from classical statistics and machine learning that are applicable to identifying multi-omic relationships with phenotypes of interest (Subramanian et al., 2020). Athieniti and Spyrou (2023) described several challenges that need to be addressed when multi-omics data analysis methods are applied. These include the ability to handle complex, non-linear relationships between features and phenotypes; missing data at one or more omics across individuals; heterogeneity in data structure because of count-based, intensity-based, or binary datasets; potential overfitting due to more variables than samples; and limited computational performance. Common methods for multi-omics data analysis include joint dimensionality reduction by investigating the covariance structure or by performing factor analysis, such as in mixOmics/DIABLO (Rohart et al., 2017; Singh et al., 2019) or MOFA (Argelaguet et al., 2018), kernel-based methods for investigating similarities; network-based methods, for instance for constructing gene regulatory networks; and deep learning methods (Athieniti & Spyrou, 2023).

1.4.2 From bulk-level to single-cell multi-omics

When multi-omics approaches are applied to characterize human diseases, the investigated biomaterials are often tissue samples from patients or from postmortem donors. While the genome does not vary between different cells and tissues, with minor exceptions, other omics such as the epigenome, the transcriptome, or the proteome are highly cell type-specific. As tissue samples represent a complex assembly of different cell types and surrounding connective tissue, the observed epigenome, transcriptome, or proteome represents an

averaged view that is dependent on the composition of the tissue. Such bulk-level multi-omics approaches can provide important insights into molecular patterns that are associated with diseases. Still, in most cases, they cannot provide information about the role of different cell types in disease. In contrast, single-cell multi-omics methods enable the profiling of multiple omics in single cells, allowing disease-associated molecular changes to be assigned to specific cell types (Baysoy et al., 2023). Single-cell RNA sequencing was the first method that enabled omics-wide analyses in single cells (Tang et al., 2009), which has led to unparalleled insights into human diseases. Today, combined profiling of the transcriptome and the chromatin architecture in single cells, for instance, using the 10X Multiome technique, allows for single-cell multi-omics approaches that can even provide insights into the regulatory mechanisms underlying the transcriptional changes (Zillich et al., 2025). Single-cell multi-omics analyses have their own methodological difficulties (Bonev et al., 2024), for example, due to data sparsity and the need for even larger computational resources. This requires adaptations of existing analysis methods for their application in the single-cell field, as realized, for instance, in MOFA+ (Argelaguet et al., 2020) or hdWGCNA (Morabito et al., 2023). While computational method development resulted in novel bioinformatic approaches, novel wet lab methods for single-cell preparation were developed, which now allow the profiling of additional omics such as the epigenome (Nichols et al., 2025) or the proteome (Ctortecka et al., 2024) in single cells. Advancing multi-omics studies to the single-cell level will thus contribute to an even better understanding of disease mechanisms in the future and enable better therapeutic options to be developed based on this knowledge.

1.4.3 Multi-omics to mechanisms

One of the central aims of multi-omics studies is to combine the entirety of genetic, epigenetic, transcriptomic, and proteomic information from a cell or tissue to derive hypotheses on disease mechanisms, which can then be investigated in functional validation experiments. Multi-omics approaches have been successfully applied to unravel disease mechanisms in various diseases such as cancers (Kan et al., 2025), cardiovascular diseases (Amrute et al., 2024), and COVID-19 (Su et al., 2020) which has even led to the identification of novel drug targets (Dimitrakopoulos et al., 2021; Rasooly et al., 2025). While multi-omics approaches were extensively applied in the cancer research field, an increasing number of studies performed multi-omics analyses in neurological and psychiatric disorders, most prominently in neurodegenerative disorders such as Alzheimer's or Parkinson's disease (Athieniti & Spyrou, 2023; Kodam et al., 2023). In the field of SUD research, several genome-, epigenome-, transcriptome- and proteome-wide association studies have been performed, for instance, in opioid use disorder (OUD) (Johnson et al., 2024). However, so far, only a few studies have performed integrative analyses of multiple omics in SUD phenotypes. Bulk-level multi-omics

approaches include an integrative analysis of DNA methylation and RNA-seq data from postmortem human brains (N=99) of the ventral and dorsal striatum in AUD, where immunological alterations were identified (Zillich, Poisel, et al., 2022). Integration of RNA-seq and proteomics data was performed in a study on opioid use disorder (OUD) identifying pro-angiogenic gene networks in postmortem PFC (N=47) (Mendez et al., 2021). Another study on OUD performed integrative analysis of GWAS data with transcriptomic, proteomic, and epigenomic data from the PFC, resulting in the identification of Akt, BDNF, and ERK signaling as conserved deregulated molecular processes in the OUD brain. In the field of single-cell multi-omics in SUDs, we have recently performed snRNA-seq and snATAC-seq in CUD for an in-parallel investigation of the transcriptome and the chromatin background in single cells of the CN in N=14 individuals (Zillich et al., 2025). We found differential expression and differential chromatin accessibility related to synapse organization and ion channels while identifying ZEB1 as a key regulator of differential expression by constructing gene regulatory networks. Another recent integrative analysis of snRNA-seq and snATAC-seq data from CN was performed in a cohort of N=143 individuals in AUD, identifying cell type-specific transcriptional and chromatin changes related to inflammation and neurotransmission (Green et al., 2024).

While the integrative analysis of multiple omics enables the identification of within-omics and between-omics relationships, it remains a key challenge to translate these findings into novel therapeutic approaches. Even if regulatory relationships are uncovered using multi-omics studies to characterize disease mechanisms, a careful and informed selection of therapeutic targets is required to develop new drugs that improve symptoms and the disease-associated phenotype. For this, validation pipelines are required to select and prioritize results for further functional investigation, for instance, in disease models. Following candidate selection according to categories such as clear linkage to disease, safety, and feasibility (Emmerich et al., 2021), functional validation in animal models reflecting addiction-like criteria such as the 3-crit model of CUD (Deroche-Gamonet et al., 2004), and drug repurposing or drug development pipelines depict important milestones on the path towards novel, mechanism-based interventions in CUD.

1.5 Aims of the studies

The following studies aimed for an in-depth characterization of molecular brain alterations in CUD, thereby providing unprecedented insights into its molecular architecture. In line with this overarching aim, several sub-objectives were pursued. One objective was to extend previous literature on individual epigenome-wide, transcriptome-wide, and proteome-wide studies in CUD using postmortem

human brain tissue by profiling multiple omics from the same individuals in multiple brain regions. Further, comparative and integrative analyses with publicly available datasets from both human CUD and rodent models aimed for the identification of reliable and conserved molecular signatures of CUD. Finally, by incorporating genetic, epigenetic, transcriptional, splicing, and protein information into a multi-omics analysis approach, the aim was to obtain a particularly comprehensive molecular characterization of brain changes in CUD to lay the foundation for the identification of disease mechanisms and potential novel therapeutic targets.

2 STUDY 1 – DNA METHYLATION IN COCAINE USE DISORDER - AN EPIGENOME-WIDE APPROACH IN THE HUMAN PREFRONTAL CORTEX[1]

2.1 Abstract

Cocaine use disorder (CUD) is characterized by a loss of control over cocaine intake and is associated with structural, functional, and molecular alterations in the human brain. At the molecular level, epigenetic alterations are hypothesized to contribute to the higher-level functional and structural brain changes observed in CUD. Most evidence of cocaine-associated epigenetic changes comes from animal studies while only a few studies have been performed using human tissue. We investigated epigenome-wide DNA methylation (DNAm) signatures of CUD in human post-mortem brain tissue of Brodmann area 9 (BA9). A total of $N = 42$ BA9 brain samples were obtained from $N = 21$ individuals with CUD and $N = 21$ individuals without a CUD diagnosis. We performed an epigenome-wide association study (EWAS) and analyzed CUD-associated differentially methylated regions (DMRs). To assess the functional role of CUD-associated differential methylation, we performed Gene Ontology (GO) enrichment analyses and characterized co- methylation networks using a weighted correlation network analysis. We further investigated epigenetic age in CUD using epigenetic clocks for the assessment of biological age. While no cytosine-phosphate-guanine (CpG) site was associated with CUD at epigenome-wide significance in BA9, we detected a total of 20 CUD-associated DMRs. After annotation of DMRs to genes, we identified *Neuropeptide FF Receptor 2 (NPFFR2)* and *Kalirin RhoGEF Kinase (KALRN)* for which a previous role in the behavioral response to cocaine in rodents is known. Three of the four identified CUD-associated co-methylation modules were functionally related to neurotransmission and neuroplasticity. Protein-protein interaction (PPI) networks derived from module hub genes revealed several addiction-related genes as highly connected nodes such as *Calcium Voltage-Gated Channel Subunit Alpha1 C (CACNA1C)*, *Nuclear Receptor Subfamily 3 Group C Member 1 (NR3C1), and Jun Proto-Oncogene, AP-1 Transcription Factor Subunit (JUN). In BA9, we observed a trend toward epigenetic age acceleration (EAA) in individuals with CUD remaining stable even after adjustment for covariates.* Results from our study highlight that CUD is associated with epigenome-wide differences in DNAm levels in BA9 particularly related to synaptic signaling and neuroplasticity. This supports findings from previous studies that report on the strong impact of cocaine on neurocircuits in the

[1] Publication: Poisel, E., Zillich, L., Streit, F., Frank, J., Friske, M. M., Foo, J. C., Mechawar, N., Turecki, G., Hansson, A. C., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2023). DNA methylation in cocaine use disorder–An epigenome-wide approach in the human prefrontal cortex. *Frontiers in Psychiatry*, *14*. https://doi.org/10.3389/fpsyt.2023.1075250

human prefrontal cortex (PFC). Further studies are needed to follow up on the role of epigenetic alterations in CUD focusing on the integration of epigenetic signatures with transcriptomic and proteomic data.

2.2 Introduction

Cocaine is one of the most frequently consumed psychostimulants in the world with around 21 million individuals having used cocaine in 2020 (UNODC, 2022). During the last decade, a stable increase in cocaine consumption and cocaine-associated hospitalization rates has been observed (European Monitoring Centre for Drugs and Drug Addiction, 2022; UNODC, 2022). While a substantial fraction of individuals retain control over cocaine intake, up to 21% of regular cocaine users develop cocaine dependence characterized by repeated cycles of binge intoxication, withdrawal, and compulsive cocaine-seeking (Lopez-Quintero et al., 2011). Chronic use of cocaine, especially crack cocaine, increases the risk for the development of cardiovascular, cerebrovascular, and respiratory diseases (Butler et al., 2017). Also, it is associated with substance abuse of other drugs such as alcohol, cannabis, and opioids (Bierut et al., 2008). Thus, chronic cocaine consumption as observed in cocaine use disorder (CUD) enhances the overall disease burden, at the individual level significantly reduces life quality, and thereby contributes to an increased risk for premature death. Investigating the pathophysiology of CUD enables the characterization of the underlying disease processes and at the same time provides a basis for the development of novel therapeutic approaches.

Substance use disorders (SUDs) including CUD are understood as brain disorders with profound alterations at the structural, functional, and molecular levels in the human brain (Koob & Volkow, 2010; Volkow et al., 2003; Volkow et al., 2019). Drug- induced effects on epigenetics are hypothesized to contribute to structural and functional changes in the brain via the establishment of altered transcriptional programs [for review see (Nestler & Lüscher, 2019; Nielsen et al., 2012; Stewart et al., 2021)]. Thus, profiling of epigenetic alterations contributes to a better understanding of molecular mechanisms underlying the neuroplastic changes associated with CUD. DNA methylation (DNAm) represents the most prominent epigenetic mechanism involved in gene expression regulation and DNAm changes in the brain following acute and chronic cocaine exposure are well described (Vaillancourt et al., 2017). So far, the effects of cocaine on DNAm were most extensively studied in the rodent brain, which has led to the identification of cocaine- associated differentially methylated genes. Interestingly, several of these genes were previously related to the neurobiology of addiction, such as Bdnf (Tian et al., 2016) and Pp1c (Anier et al., 2010) as well as the immediate early genes c-Fos (Wright et al., 2015) and fosB (Anier et al., 2010). Cocaine was also shown to affect methylation and expression levels of DNA methyltransferases (Dnmt1, Dnmt3a) (Anier et al., 2013; Cannella et al., 2018; Tian et al., 2012; Wright et al.,

2015) and Tet demethylases (Feng et al., 2015) thereby directly altering DNAm homeostasis.

In contrast, much less is known about cocaine-associated DNAm changes in humans. Due to an easy and minimally invasive sampling procedure, epigenome-wide association studies (EWASs) of different cocaine use phenotypes have been performed in human blood samples (Camilo et al., 2019; Shu et al., 2022). These studies revealed an enrichment of cocaine-associated differential methylation within biological pathways involved in synaptic transmission, neurogenesis, cell division, and chromatin organization. However, as DNAm levels are known to be tissue-specific, the predictive value of differential methylation in blood for methylation signatures of CUD in the brain might be limited. Thus, profiling DNAm levels in brain tissue is of critical importance. So far, two epigenome-wide characterizations in CUD have been performed in human post-mortem brain tissue from the caudate nucleus (CN) (Vaillancourt, Yang, et al., 2021) and the nucleus accumbens (NAc) (Vaillancourt, Chen, et al., 2021). By analyzing post-mortem tissue of the CN from N = 25 CUD cases and N = 25 controls using a reduced representation bisulfite sequencing approach, a total of 173 differentially methylated regions (DMRs) were found that were significantly associated with CUD status. The IRX2 gene was among the top findings of CUD-associated differential methylation in the CN. Using a similar approach in the NAc, the key finding was a hypermethylated region within the tyrosine hydroxylase (TH) gene encoding an enzyme involved in dopamine metabolism. Interestingly, the identified DMR in TH was shared between the CN and the NAc. Both studies provide some first mechanistic insights into the biological relevance of CUD-associated DNAm changes in the human brain.

While epigenome-wide alterations associated with CUD have been studied in the human striatum, they have not been reported for cortical brain regions so far. The prefrontal cortex (PFC) has a key role in the addiction cycle due to its relation to dysregulated cognitive processes in SUDs such as inhibitory control and response to drug- related cues (Goldstein & Volkow, 2011). In CUD, resting-state connectivity alterations have been found in frontostriatal circuits representing functional connections between the PFC and the striatum (Y. Hu et al., 2015). As DNAm alterations in the PFC have been described for the Brodmann area 9 (BA9) subregion in other SUDs such as alcohol use disorder (AUD) (Lohoff et al., 2021; Zillich, Frank, et al., 2022) and opioid use disorder (OUD) (Liu et al., 2021; Shu et al., 2021), BA9 depicts a promising brain region for the epigenome-wide profiling of DNAm in CUD.

Based on epigenome-wide DNAm data, epigenetic clocks provide the opportunity to estimate biological age, a measure that increases naturally with chronological age, but is strongly influenced by disease processes such as chronic inflammation

(Levine et al., 2018). Two commonly used methodologies are Horvath's (Horvath, 2013) and Levine et al. (2018) (PhenoAge) epigenetic clocks. Epigenetic age acceleration (EAA), defined as the difference between the estimated epigenetic and chronological age, is of particular interest in the context of diseases, with its interpretation as faster or slower aging considering chronological age as the reference. Existing literature reports on a trend toward positive EAA in the PFC of individuals with SUDs such as AUD and OUD compared to a control population, especially for Levine's clock (Cabrera-Mendoza et al., 2023; Shu et al., 2021). This implies that the SUD brain appears to be epigenetically older compared to a brain from an individual without SUD at the same chronological age.

In the present study of CUD in human post-mortem brain tissue of BA9, we aimed to expand the epigenome-wide characterization of CUD to the PFC as another key brain region in addiction. Next to performing an EWAS of CUD in BA9, we aimed to identify functional enrichment of CUD-associated differential methylation within biological pathways and utilized epigenetic clocks to investigate the relationship between chronological and biological age in CUD.

2.3 Materials & Methods

2.3.1 Post-mortem human brain samples

Post-mortem human brain tissue of BA9 from $N = 42$ male individuals was received from the Douglas Bell Canada Brain Bank (DBCBB, Montréal, QC, Canada). Brain samples were obtained as fresh-frozen tissue and were stored at −80°C until processing. All sample donors were of European ancestry. At the brain bank, phenotypic information was routinely obtained using a psychological autopsy including next-of-kin interviews complemented by the coroner's notes and the medical record of the deceased donor. Psychiatric diagnoses related to substance use and comorbidities were based on the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV classification system, however, we used the more recent nomenclature of DSM-5, i.e., SUD instead of substance dependence throughout this study. We included subjects with donor age > 18 and information on CUD status (yes = CUD, no = no CUD). Exclusion criteria were severe neurodevelopmental and neuropsychiatric disorders, SUDs other than CUD or AUD, and psychiatric diagnoses except for major depressive disorder diagnoses (MDD, depressive disorder NOS). Matching was performed by age and comorbidities resulting in $N = 21$ CUD case/control pairs.

2.3.2 DNA extraction and methylation data analysis

Extraction of DNA from $N = 42$ BA9 post-mortem brain samples was performed using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). DNA was

stored at −20°C and epigenome-wide DNAm was determined using the Illumina Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA). Samples were randomized based on CUD case/control status and comorbidities such as depressive disorders and AUD prior to the array processing steps.

Quality control (QC) was performed on raw data using a modified version of the CPACOR-pipeline (Lehne et al., 2015) as described in Zillich, Frank, et al. (2022). All statistical analyses were performed using the statistical computing environment R v4.2.1 (R Core Team, 2022). In brief, sample exclusion criteria were missing rate > 0.10 and discordance between phenotypic and predicted sex based on DNAm. cytosine-phosphate- guanine (CpG) sites were excluded if (1) the call rate was below 0.95, (2) they are located on sex chromosomes, and (3) their genomic position is related to a genetic variant with MAF > 0.10.

Raw intensities were quantile-normalized followed by a conversion to beta-values of methylation. Beta-values were logit- transformed to *M*-values as recommended by Du et al. (2010). To assess DNAm signatures of CUD in post-mortem brain tissue, we applied a linear regression model using *M*-values as the dependent variable and CUD status as the predictor. Donor age, post-mortem interval (PMI), brain pH, diagnosis of a depressive disorder, and AUD status were included as covariates in the regression model. We performed a variance partition analysis using the R package variancePartition v1.26.0 (Hoffman & Schadt, 2016) to confirm the covariates (see also Supplementary Figure 2.1). As variance explanation by cause of death was only minimal and would have made the statistical model significantly more complex, we did not include it as a covariate. Based on the methylation data, we estimated the fraction of neuronal and non-neuronal cells with the Houseman (Houseman et al., 2012) approach using the dorsolateral PFC (dlPFC) dataset (Jaffe & Kaminsky, 2017) as reference. We included the fraction of neuronal cells as a covariate to correct for cell-type heterogeneity in the samples. To adjust for technical effects from microarray processing, we derived principal components from the internal control probes of the EPIC array and included the first 10 principal components (cpPC1–10) into the statistical model.

This resulted in the following linear model:

$$DNAm \sim CUD\ status + age + PMI + pH + depressive\ disorder\ status + AUD\ status + neuronal\ fraction + cpPC1 + cpPC2 \ldots + cpPC10 \quad (1)$$

We performed Benjamini–Hochberg false discovery rate (FDR) correction (Benjamini & Hochberg, 1995) to adjust for multiple testing. Annotation of CpG

sites to genes and genomic background was based on the Illumina manifest as downloaded on 10th August 2018:

http://webdata.illumina.com.s3- website- us- east- 1.amazonaws. com/downloads/productfiles/methylationEPIC/infinium- methylationepic- v- 1- 0- b4- manifest- file- csv.zip.

To test for enrichment of CUD-associated CpG sites within categories of gene regulatory features (TSS1500, TSS200, 5′- UTR, 1st Exon, ExonBnd, Body, 3′-UTR, and intergenic regions) and the CpG island (CGI) background (N-Shelf, N-Shore, Island, S-Shore, S-Shelf, and open sea regions), we performed chi-squared tests against the distribution of features among all probes on the EPIC array. The results of the enrichment analysis were adjusted for multiple testing using Bonferroni correction.

2.3.3 Differentially methylated regions

Following the assessment of CUD-associated methylation differences at single CpG sites, we investigated DMRs associated with CUD status. We used the dmrff algorithm v1.0.0 (Suderman et al., 2018) for the identification of DMRs which was recently shown to outperform several other DMR identification tools when applied to methylation array data (Lent et al., 2021). We applied dmrff on the EWAS summary statistics (N = 654,448 CpG sites). The maximum distance between individual CpG sites (maxgap) was set to 500 bp and an association p-value threshold (p. cutoff) of 0.01 was used. The implemented Bonferroni correction was used to adjust for multiple testing during the identification of DMRs.

2.3.4 Replication of CpG-sites identified in the striatum and in whole blood

We further tested whether DNAm signatures, which had previously been associated with CUD in whole blood (Camilo et al., 2019), the CN (Vaillancourt, Yang, et al., 2021), and the NAc (Vaillancourt, Chen, et al., 2021) were also associated with CUD in the present sample of BA9. For DMRs overlapping between the two studies in the striatum, we tested CpG sites in close proximity (500 bp up- and downstream) for their association with CUD in our sample. The comparison with CUD-associated CpG sites derived from whole blood was restricted to the subset of CpG sites that was available in our analysis in BA9.

2.3.5 Gene-Ontology enrichment analysis

To investigate the potential biological relevance of CUD-associated differential methylation, we analyzed the overrepresentation of differentially methylated CpG sites within biological pathways. Under the assumption of hypomethylation leading to increased gene expression and vice versa, we applied missMethyl

v1.30.0 (Phipson et al., 2016) separately to hypermethylated and hypomethylated CpG sites after selection for a CUD association p-value < 0.001. The set of N = 654,448 CpG sites remaining after QC was used as the reference set during the enrichment analysis. Further, we performed an enrichment analysis on DMRs resulting from dmrff using the "goregion" function in missMethyl. Correction for multiple testing in results from missMethyl was performed using FDR.

2.3.6 WGCNA

To further characterize the biological and functional relevance of CUD-associated methylation signatures in the brain, we performed a weighted correlation network analysis using WGCNA v1.71 (Langfelder & Horvath, 2008). WGCNA identifies co-methylation modules and provides a correlation estimate between modules and traits such as CUD status. We prioritized CpG sites based on their annotation to promoter/transcription start sites (TSS 200 and TSS1500) under the assumption of a *cis*-regulatory role of their methylation levels on gene expression. Quantile normalized beta values of the resulting 117,876 CpG sites were used as input data for WGCNA. Soft power threshold picking was performed according to the R_{signed}^2 > 0.90 criterion resulting in a power threshold of five for the methylation data in BA9. Parameters for constructing the signed network in a block-wise approach were minModuleSize = 30, maxBlockSize = 36,000, and mergeCutHeight = 0.25. For the CUD-associated co-methylation modules, enrichment analysis was performed in missMethyl using the N = 117,876 input CpG sites as background. Module hub genes were derived by annotation of CpG sites to genes and were defined as the top 0.5% of annotated genes in a co-methylation module ranked by the product of module membership (correlation of module eigengene with methylation profile) and gene significance (correlation between CpG site and trait). Gene significance was calculated for each CpG site. Protein-protein interaction (PPI) networks of hub genes from the CUD-associated co-methylation modules were generated in Cytoscape v3.9.1 (Shannon et al., 2003) using the stringApp v1.7.0 (Doncheva et al., 2019). Networks of the "full STRING network" type were generated by applying the STRING protein query option on the hub gene list. An interaction score threshold of 0.7 was used to include only high-confidence interactions. PPI networks were visualized using the "Prefuse Force Directed Layout" option while unconnected nodes ("Singletons") and hidden edges were removed from the visualization. Information on stringdb identifiers corresponding to the nodes of the final PPI network for each module is provided in Supplementary Tables 2.7A–D.

2.3.7 Epigenetic age estimates

Using quantile-normalized beta-values of the BA9 methylation data, we estimated EAA using epigenetic clocks as implemented in the R package methylclock v1.2.1

(Pelegí-Sisó et al., 2021). We applied Horvath's (Horvath, 2013) (Horvath) and Levines's (Levine et al., 2018) (PhenoAge) epigenetic clocks to assess epigenetic age in the context of CUD directly in the brain. We first investigated the relationship between epigenetic age and chronological age using correlation analysis. Next, we estimated EAA (ageAccel), defined as the residuals of regressing chronological age on epigenetic age. We performed *t*-tests to assess the group difference between individuals with and without CUD and linear regression models to evaluate the association with CUD status while adjusting for the covariates neuronal cell fraction, pH, PMI, depressive disorder, and AUD status.

2.3.8 Ethics approval statement

Post-mortem brain tissue sampling at the DBCBB was performed according to their established ethical standards including written informed consent from the next-of-kin for each subject. Our study design was approved by the Ethics Committee II of the University of Heidelberg, Medical Faculty Mannheim, Germany, under the register number 2021-681.

2.4 Results

Demographic data of sample donors are summarized in Table 2.1, while a detailed overview of toxicology and cause of death is provided in Supplementary Table 2.1. After QC of methylation data, all $N = 42$ subjects, and a total of 654,448 CpG sites remained for further analysis.

Table 2.1 - Demographic data of sample donors included in the present study on CUD
SD, standard deviation; PMI, post-mortem interval; MDD, major depressive disorder; NOS, depressive disorder not otherwise specified; *p*-value, derived from CUD/no CUD comparison using a *t*-test for continuous and chi-squared test for categorical variables.

Characteristic	CUD	No CUD	p-value
N	21	21	-
Age, years (SD)	44.4 (13.0)	50.1 (13.8)	0.17
Sex (Male/Female)	21/0	21/0	-
pH (SD)	6.37 (0.39)	6.26 (0.28)	0.33
PMI, hours (SD)	56.8 (23.9)	58.4 (19.9)	0.82
Cocaine or cocaine metabolites at death, yes (%)	15 (71.4)	0 (0)	-
Depressive disorder (MDD, NOS), yes (%)	9 (42.9)	5 (23.8)	0.33
Alcohol use disorder, yes (%)	9 (42.9)	4 (19.1)	0.18

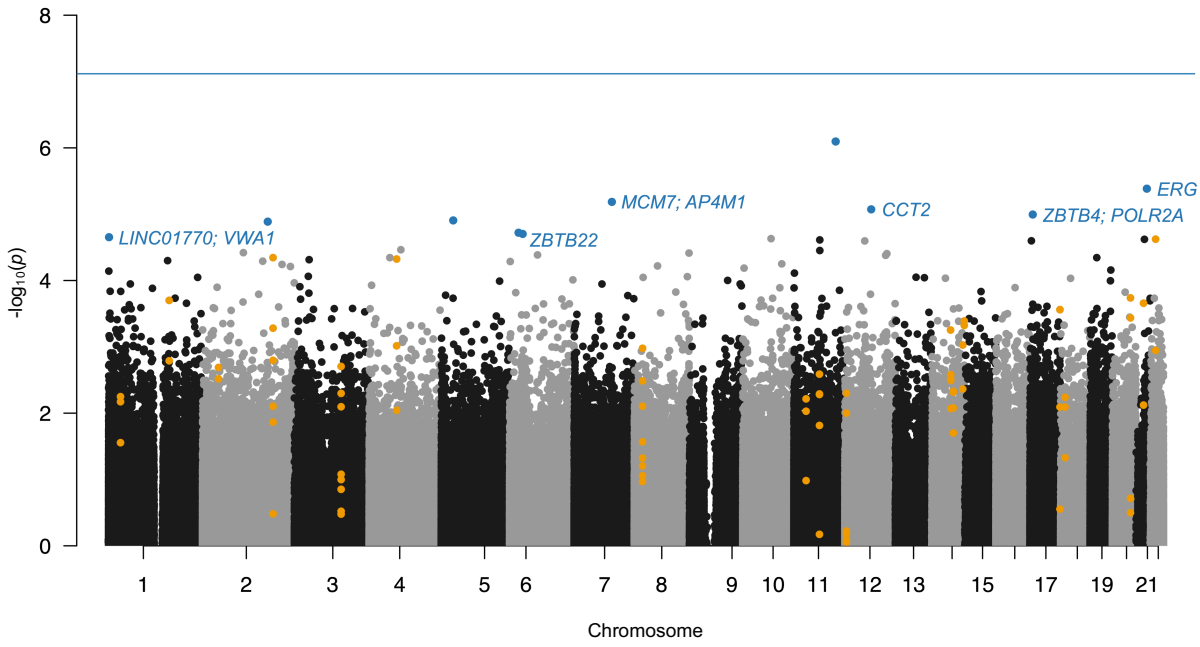
2.4.1 Epigenome-wide association study of CUD

None of the CpG sites passed epigenome-wide significance ($p < 7.64e{-}08$). The strongest association with CUD in BA9 was detected for the CpG site cg04659218 ($p = 7.98e{-}07$, FDR = 0.52) on chromosome 11. A Manhattan plot summarizing the results is displayed in Figure 2.1A. Descriptively, several of the CpG sites showing stronger associations with CUD status were annotated to genes encoding transcription factors and DNA binding domain- containing proteins: *ETS Transcription Factor ERG* (*ERG*, $p = 4.15e{-}06$), *Zinc Finger and BTB Domain Containing 4* (*ZBTB4*, $p = 1.01e{-}05$), *RNA Polymerase II (POL-II) Subunit A* (*POLR2A*, $p = 1.01e{-}05$), and *Zinc Finger and BTB Domain Containing 22* (*ZBTB22*, $p = 1.97e{-}05$, see Figure 2.1B). EWAS summary statistics of the 1,000 CpG sites displaying the strongest association with CUD status are provided in Supplementary Table 2.2. A Q–Q plot ($\lambda = 0.969$) derived from the EWAS *p*-values is shown in Supplementary Figure 2.2.

2.4.2 Differentially methylated regions associated with CUD status

A total of 20 DMRs were significantly associated with CUD status after multiple testing correction. We detected more hypermethylated ($N = 17$, 85%) than hypomethylated ($N = 3$, 15%) DMRs. The DMR consisting of six hypermethylated CpG sites annotated to the genes *LOC101927196* and *Fibrous Sheath Interacting Protein 2 (FSIP2)* displayed the strongest association with CUD status in BA9 ($p = 1.74e{-}16$). Results of the DMR analysis including summary statistics for all 20 identified regions are visualized in Table 2.2.

A



B

	chr	pos	cg	Gene	BETA	SE	P_VAL	N	FDR
1	11	109827002	cg04659218		-0,436	0,066	7,98e-07	42	0,522
2	21	39796367	cg12281271	*ERG*	0,521	0,088	4,15e-06	42	0,999
3	7	99701060	cg04747451	*MCM7; AP4M1*	-0,432	0,075	6,54e-06	42	0,999
4	12	69982777	cg12766007	*CCT2*	-0,461	0,082	8,47e-06	42	0,999
5	17	7387573	cg03443505	*ZBTB4; POLR2A*	-0,840	0,151	1,01e-05	42	0,999
6	5	32204300	cg19102526		-0,383	0,070	1,24e-05	42	0,999
7	2	172372096	cg02097254		-0,568	0,104	1,31e-05	42	0,999
8	6	25218891	cg02224516		0,605	0,114	1,92e-05	42	0,999
9	6	33285470	cg20225889	*ZBTB22*	-1,003	0,189	1,97e-05	42	0,999
10	1	1369793	cg00702126	*LINC01770; VWA1*	0,298	0,057	2,22e-05	42	0,999

Figure 2.1 - Results of the EWAS of CUD assessing $N = 654{,}448$ cytosine-phosphate-guanine (CpG) sites in $N = 42$ post-mortem human brain samples from Brodmann area 9 (BA9). **A** In the Manhattan plot, the 10 CpG sites displaying the strongest association with CUD status were highlighted in blue and were annotated to gene name if available. The 20 CUD-associated DMRs are highlighted in orange. The solid blue line depicts epigenome-wide significance ($p < 7.64e{-}08$). **B** EWAS summary statistics for the 10 CpG sites showing the strongest association with CUD status.

2.4.3 Replication of CpG sites identified in whole blood and striatal post-mortem brain tissue

Of the 186 CpG sites associated with cocaine and crack cocaine use in whole blood, 138 were available in our dataset after QC. Six CpG sites annotated to the genes *PDSS2*, *EHMT2*, *C2*, *PDE12*, *ARF4*, *ARF4-AS1*, *RMDN2*, *RMDN2-AS1*, and *CDKN1C* were nominally associated in the present analysis (see also Supplementary Table 2.3A) but did not remain significant after multiple testing correction. The DMRs previously identified as being conserved between the CN and NAc annotated to the genes *SPEG*, *TH*, *PDXDC1*, *LDHD*, and *SYT5* were not

replicated in our sample of BA9. Similarly, CpG sites located in a genomic window 500 bp up- and downstream of these overlapping striatal DMRs were not significantly associated with CUD in BA9 (Supplementary Table 2.3B). However, when we compared the DMRs in BA9 with DMRs detected in the CN and the NAc individually, we found one hypermethylated CUD- associated DMR annotated to the *Aldehyde Dehydrogenase 3 Family Member B1 (ALDH3B1)* gene on chromosome 11 that was conserved between BA9 and the CN.

Table 2.2 - CUD-associated differentially methylated regions (N = 20). Differentially methylated regions associated with CUD status.
Pos, position of the most significant CpG site within the DMR, BETA, and P-VAL: effect size and *p*-value of this CpG site derived from the epigenome-wide association study (EWAS), N CpG, number of CpG sites within DMR, DMR *z*, and DMR *q*-val: DMR statistics resulting from *dmrff* (Suderman et al., 2018).

	Chr	Pos	Gene	BETA	P VAL	DMR start	DMR end	N CpG	DMR z	DMR q-val
1	*2*	186603233	*LOC101927196; FSIP2*	0.28	4.52e-05	186603233	186603441	6	9.71	1.74e-16
2	*14*	76734493		0.63	4.69e-03	76734327	76734550	4	8.79	9.54e-13
3	*17*	79339679		0.39	2.75e-04	79339278	79339679	3	7.69	9.79e-09
4	*11*	67777715	*ALDH3B1*	0.22	2.58e-03	67777682	67777952	5	7.58	2.25e-08
5	*4*	72897782	*NPFFR2*	2.69	4.74e-05	72897613	72897782	3	7.46	5.74e-08
6	*18*	11751352	*GNAL*	0.24	5.76e-03	11751317	11751352	3	6.99	1.84e-06
7	*3*	123813596	*KALRN*	0.33	1.98e-03	123813383	123813596	8	6.95	2.40e-06
8	*8*	22132992	*LOC100507071; PIWIL2*	0.72	1.05e-03	22132641	22132992	8	6.62	2.38e-05
9	*21*	30449989	*MAP3K7CL*	0.47	2.20e-04	30449989	30450130	2	6.50	5.30e-05
10	*22*	30476204	*HORMAD2-AS1; HORMAD2*	0.64	2.38e-05	30476204	30476206	2	6.35	1.47e-04
11	*1*	161008297	*TSTD1*	0.41	1.99e-04	161008127	161008297	2	6.32	1.68e-04
12	*20*	47896888	*ZFAS1; SNORD12B; SNORD12*	0.34	1.83e-04	47896888	47897207	4	5.97	1.59e-03
13	*2*	42274082	*PKDCC*	0.42	2.05e-03	42274082	42274306	2	5.89	2.62e-03
14	*12*	4382184	*CCND2-AS1; CCND2*	-0.39	5.01e-03	4382114	4382184	6	-5.81	4.07e-03
15	*14*	105915268	*MTA1*	0.25	4.08e-04	105915004	105915268	2	5.74	6.17e-03
16	*14*	101908830	*LOC100507277*	-0.33	9.41e-04	101908830	101908865	2	-5.66	1.02e-02
17	*14*	70193496		-0.41	2.63e-03	70193468	70193496	3	-5.58	1.56e-02
18	*14*	69656853	*EXD2*	0.45	5.58e-04	69656853	69657279	2	5.58	1.62e-02
19	*11*	32421635	*WT1*	0.34	6.09e-03	32421514	32421635	3	5.41	4.28e-02
20	*1*	32169770	*COL16A1*	0.35	5.64e-03	32169701	32169770	3	5.38	4.89e-02

2.4.4 Functional annotation and enrichment analysis of CUD-associated CpG sites

To analyze enrichment within genomic features (Figure 2.2A) and the CGI background (Figure 2.2B), we prioritized our findings based on CpG sites associated with CUD at nominal significance ($p < 0.05$). This resulted in a subset of $N = 29{,}176$ CpG sites of which $N = 14{,}737$ were hypermethylated and $N = 14{,}439$ were hypomethylated. We detected a balanced ratio between hyper- and hypomethylation in this subset (Figure 2.2C). For the CUD-associated CpG sites,

a statistically significant enrichment within 1st Exon, N-Shore, and CGI regions was detected, whereas a significant depletion from intergenic and open sea regions was found.

The most significantly enriched Gene Ontology (GO) term was *response to exogenous dsRNA* (p = 3.71e–04) for hypermethylated CpG sites and *regulation of calcium ion transmembrane transporter activity* (p = 2.39e–04) for the subset of hypomethylated CpG sites. No enrichment remained statistically significant after multiple testing correction (Supplementary Table 2.4A).

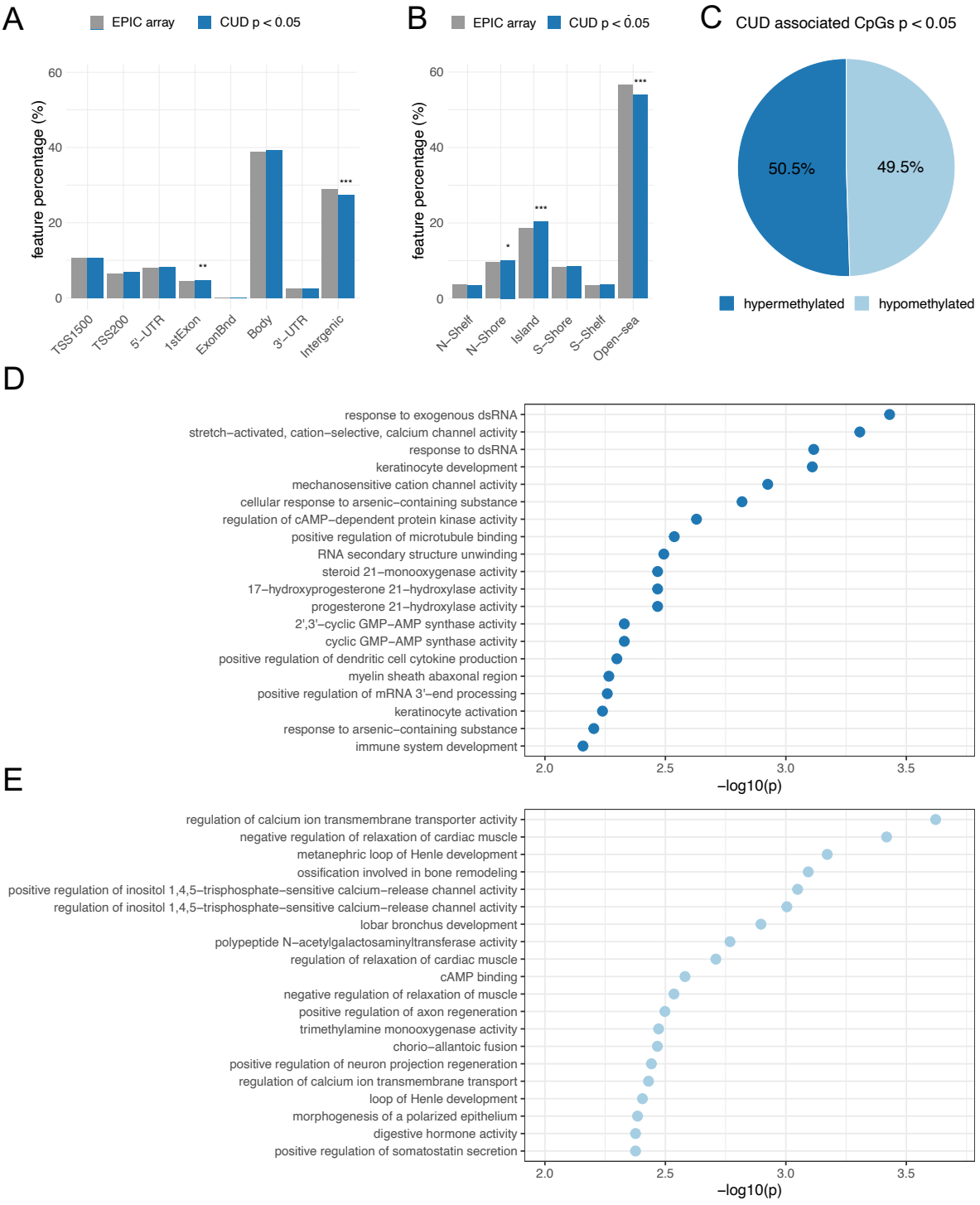


Figure 2.2 - Results from the epigenome-wide association study (EWAS) of CUD were prioritized based on association *p*-values resulting in a subset of $N = 29{,}176$ CpG sites with $passoc < 0.05$ that was further analyzed by annotation to panel **A** gene regulatory features and **B** the CpG island (CGI) background. The proportion of CUD-associated CpG sites (blue) within different groups was compared to the EPIC array background (gray). Differences remaining statistically significant after Bonferroni multiple testing correction are highlighted using asterisks (* = $padj < 0.05$, ** = $padj < 0.01$, *** = $padj < 0.001$). **C** Hypermethylation (light-blue) and hypermethylation (dark blue) among the $N = 29{,}176$ CUD-associated CpG sites. The top 20 CUD-associated GO terms resulting from missMethyl (Phipson et al., 2016) are shown for **D** hypermethylated and **E** hypomethylated CpG sites associated with CUD ($p < 0.001$).

Descriptively, genes harboring hypermethylated CpG sites were overrepresented within biological pathways involved in cation and especially mechanosensitive calcium ion channel activity, cyclic adenosine monophosphate (cAMP)-dependent protein kinase activity (PKA), and immunological signaling mainly related to the cyclic GMP-AMP synthase (CGAS)-mediated response to nucleic acids (Figure 2.2D). In contrast, the gene set derived from hypomethylated CpG sites was enriched within pathways related to calcium transmembrane transport, negative regulation of muscle relaxation, and cAMP binding (Figure 2.2E). Also, *positive regulation of inositol-1,4,5- trisphosphate-sensitive calcium release channel activity* was among the enriched GO terms detected for hypomethylated CpG sites. The strongest overrepresentation resulting from the enrichment analysis using DMRs was observed for "thiosulfate-thiol sulfurtransferase activity" (Supplementary Table 2.4B).

2.4.5 Weighted correlation network analysis

Due to the known *cis*-regulatory role of DNAm on gene expression, we extracted TSS1500- and TSS200-associated CpG sites from the methylation matrix to identify transcriptionally relevant co- methylation networks. After hierarchical clustering in WGCNA, the sample tree was pruned at the height of 20 resulting in the removal of one sample from the analysis that was considered as an outlier (N = 41 remaining). The sample dendrogram was related to traits resulting in the clustering of CUD cases and non-CUD individuals within subgroups (Figure 2.3A). Network construction led to the identification of a total of 69 modules ranging in size between 38 and 35,758 CpG sites. A module-trait correlation heatmap for all resulting modules is shown in Supplementary Figure 2.3.

Of the 69 modules, four were significantly associated with CUD status: brown (N = 6,309 CpG sites, p = 9.15e–03), brown4 (N = 187 CpG sites, p = 0.04), blue (N = 14,050 CpG sites, p = 0.02), and steelblue (N = 340 CpG sites, p = 0.02). A positive, highly significant correlation between module membership and gene significance for CUD status was detected within these modules (Figure 2.3B). Whereas module brown was positively associated with CUD status (r = 0.40), a negative correlation was identified for modules brown4, blue, and steelblue ($-0.37 \leq r \leq -0.32$, see Figure 2.3C). In addition, for the brown module, significant negative correlations were detected with age (r = −0.5, p = 8.11e–04) and the fraction of neuronal cells (r = −0.91, p = 7.69e–17). The blue module also depicts a cell type-specific network displaying a highly significant positive correlation with neuronal fraction (r = 0.96, p = 1.78e–22). To further assess the biological processes related to CUD-associated modules, we performed GO enrichment analyses (Figure 2.3D). After multiple testing correction, GO enrichment in modules brown and blue remained statistically significant. Descriptively, module brown was enriched for synaptic signaling and neurotransmitter transport, whereas module brown4 revealed GO

terms related to transcription factor binding and neuroblast as well as synapse maturation. The blue module displayed enrichment of large GO terms involved in basic cellular processes and compartments, such as extracellular region, plasma membrane, and cell adhesion. Within module steelblue, nominally significant pathway enrichment was related to ion channel activity, membrane potential homeostasis, and vesicle transport. Full enrichment statistics for the top enriched GO terms in each module are provided in Supplementary Table 2.5.

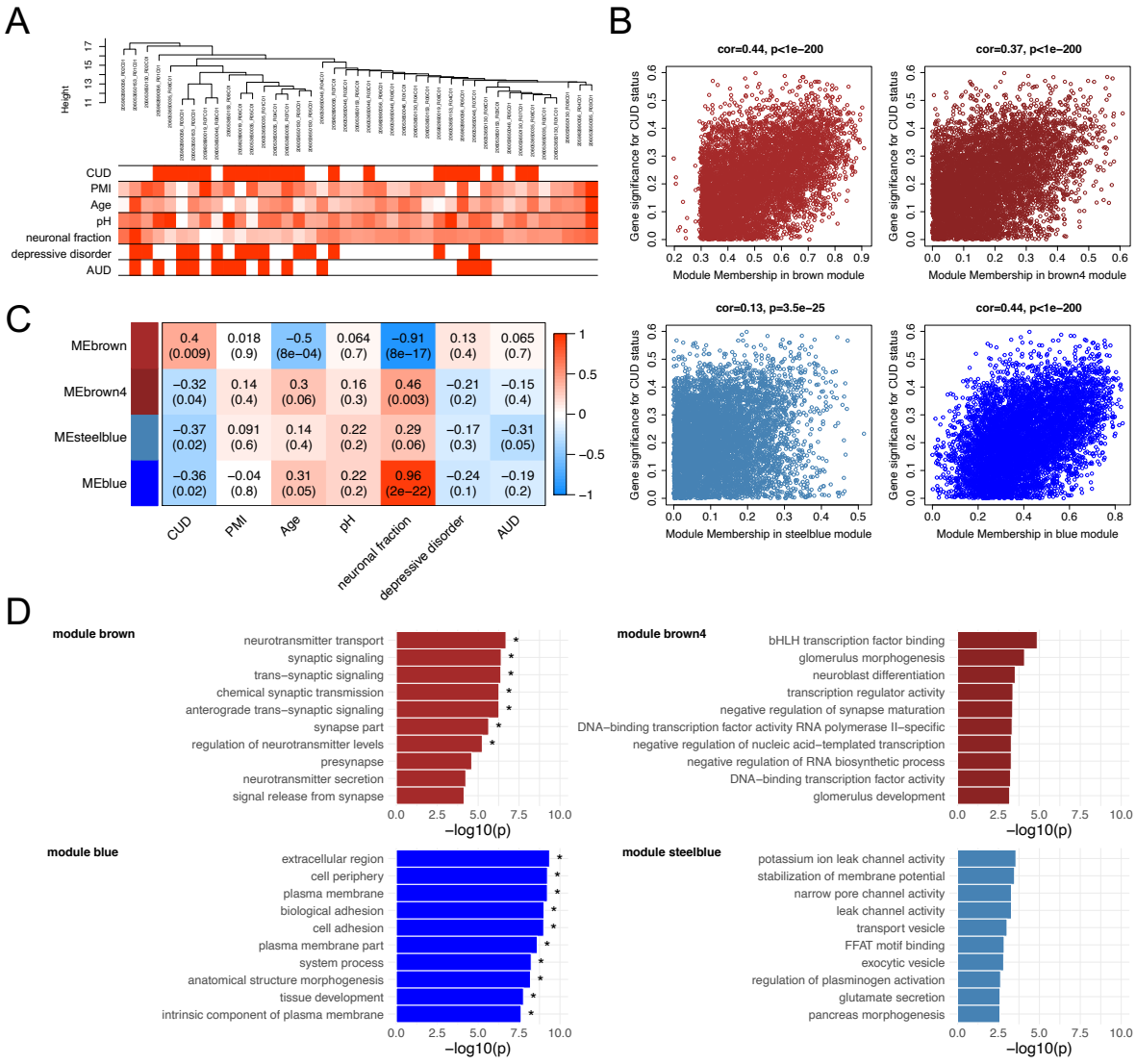


Figure 2.3 - Functional analysis of the CUD-associated co-methylation modules identified in WGCNA. Quantile-normalized beta values of methylation for N = 117,876 promoter-associated cytosine-phosphate-guanine (CpG) sites (TSS200 and TSS1500) were extracted and used as an input dataset for WGCNA. **A** Dendrogram of hierarchical sample clustering with trait annotation. **B** Correlation plots of module membership and CUD gene significance for the four CUD-associated co-methylation modules (brown, brown4, steelblue, blue). **C** Module-trait relationships for the CUD-associated modules. Module-trait correlation coefficients are color-coded, and correlation p-values are shown in brackets. **D** Results of the Gene Ontology (GO) enrichment analysis. The 10 GO terms with the strongest enrichment within the respective co-methylation module are displayed (* = FDR < 0.05). CUD, cocaine use disorder; PMI, post-mortem interval; neuronal fraction, estimated neuronal cell type proportion based on Houseman et al. (2012); AUD, alcohol use disorder.

To further characterize the molecular processes enriched in the CUD-associated modules, we constructed PPI networks based on module hub genes (listed in Supplementary Table 2.6). The resulting PPI network plots are displayed in Supplementary Figures 2.4A– D. Among PPI network hub nodes of modules brown and blue (Supplementary Table 2.8), we found members of the Transforming Growth Factor Beta (TGF-ß) signaling cascade such as *TGF Alpha* (*TGFA*), *TGFB1*, and *Endoglin* (*ENG*). Reflecting the pathway enrichment for neuron-related biological processes in modules brown, brown4, and steelblue, we detected multiple hub nodes with an important role in neuronal activity and function within these modules. Among them were the Ca*v*1.2 subunit (*CACNA1C*, module steelblue), the glucocorticoid receptor (*NR3C1*, module brown), the acetylcholine transporter *Solute Carrier Family 18 Member A3* (*SLC18A3*, module brown4), and *Vesicle Associated Membrane Protein 2* (*VAMP2*, module steelblue). Further, several tyrosine kinases and their downstream targets were found among hub nodes of PPI networks derived from modules brown4 and blue, such as *Fibroblast Growth Factor Receptor 1* (*FGFR1*), *Erb-B2 Receptor Tyrosine Kinase 3* (*ERBB3*), *KRAS Proto-Oncogene GTPase* (*KRAS*), and *Signal Transducer And Activator Of Transcription 5A* (*STAT5A*). The *JUN* was the node displaying the strongest connectivity in module steelblue.

2.4.6 Epigenetic age in CUD

A moderate (R_{Levine}^2 = 0.44) to strong ($R_{Horvath}^2$ = 0.80) correlation was found between estimated epigenetic age and chronological age (Figures 2.4A, B).

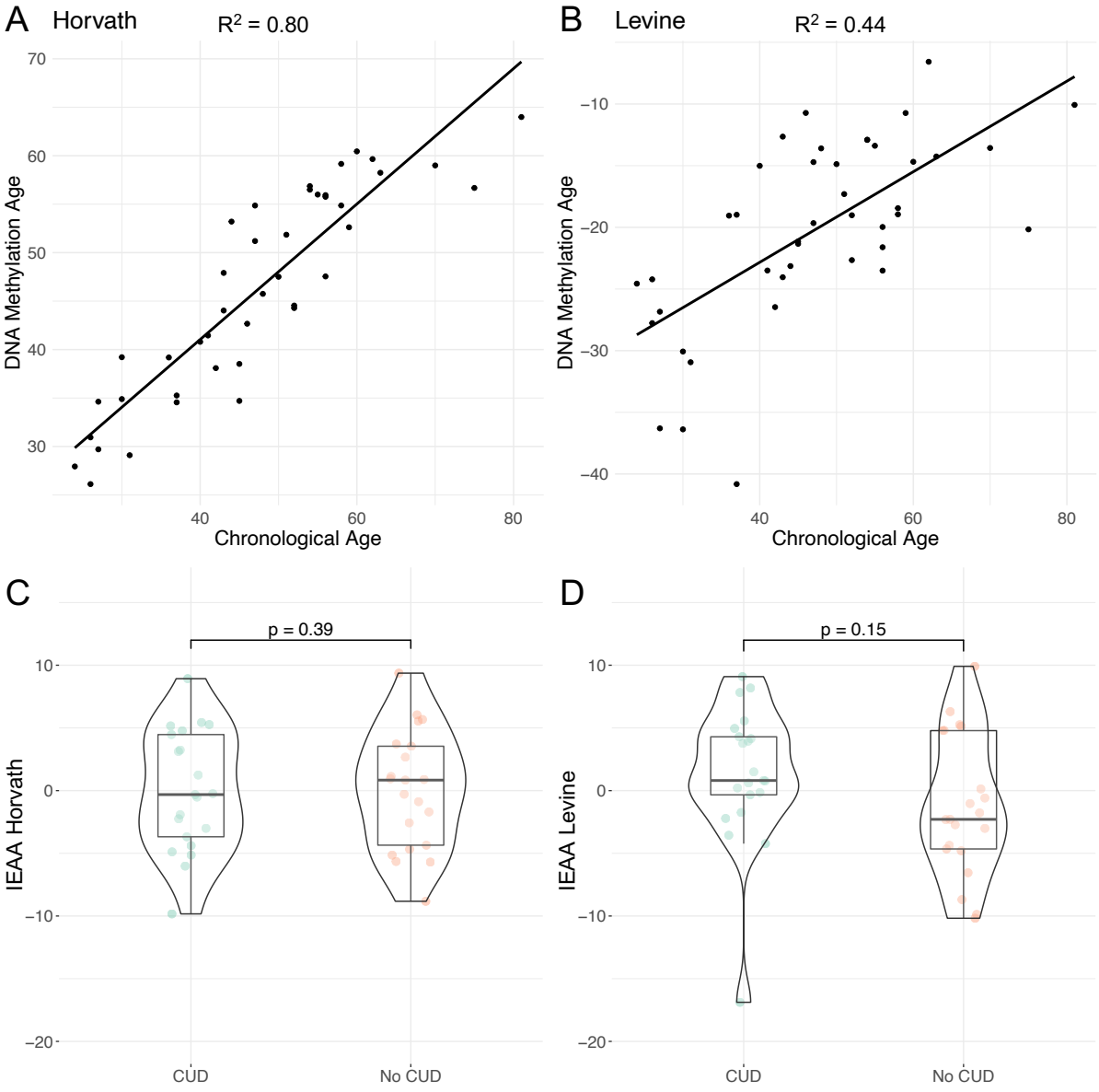


Figure 2.4 Evaluation of epigenetic clocks in cocaine use disorder (CUD) using post-mortem human brain samples of Brodmann Area 9 (BA9). Correlation of chronological age with DNA methylation (DNAm) based **A** Horvath's and **B** Levine's epigenetic age estimates as implemented in the R package methylclock (Pelegí-Sisó et al., 2021). Comparison of ageAccel between individuals with and without CUD using **C** Horvath's and **D** Levine's epigenetic clocks.

Levine's epigenetic clock produced strongly negative PhenoAge estimates (Figure 2.4B). When assessing ageAccel we observed a trend toward a more positive ageAccel in CUD cases compared to individuals without CUD, however, not statistically significant ($p_{Horvath}$ = 0.39, p_{Levine} = 0.15, Figures 2.4C, D). For both epigenetic clocks, we consistently found positive associations for CUD status

while adjusting for covariates in the regression models for ageAccel, although not statistically significant (Supplementary Table 2.9).

2.5 Discussion

We performed an epigenome-wide analysis of DNAm differences in BA9 post-mortem human brain tissue from $N = 21$ CUD cases and $N = 21$ individuals without CUD. In the BA9 subregion of the human PFC, none of the associations of CpG site methylation levels with CUD status passed the level of epigenome-wide significance. However, we identified 20 DMRs significantly associated with CUD.

Among the strongest single-site associations with CUD status, there were multiple CpG sites annotated to genes involved in transcriptional regulation either *via* intrinsic transcription factor activity or due to the interaction with RNA POL-II. The transcription factor ERG is involved in the regulation of vasculature integrity (Birdsey et al., 2015). For ZBTB4, binding to methylated CpG sites was reported (Roussel-Gervais et al., 2017) and ZBTB22 is assumed to be involved in transcription regulation *via* interaction with POL-II (Safran et al., 2021). In our EWAS, we also detected CUD-associated differential methylation within the largest subunit of POL-II (POL2RA), suggesting a role of CUD-associated methylation signatures in gene expression regulation. Results of the enrichment analysis for CUD-associated CpG sites within genomic features are in line with this: the significant enrichment of CUD-associated differential methylation within features containing transcription factor binding sites and the depletion from intergenic and non-CGI- associated (open sea) regions points toward a regulatory potential of CUD-associated DNAm patterns on gene expression.

For several of the genes harboring DMRs significantly associated with CUD in BA9, there is evidence for their relevance for neuronal function. The DMR displaying the strongest association with CUD status was related to the *LOC101927196* gene. Overexpression of the long non-coding RNA LOC101927196 was shown to reduce cell proliferation and support apoptotic processes in the brain in a rat model of autism (Yao et al., 2019). For *NPFFR2* and *KALRN*, we found direct links to their importance for the biological response to cocaine. NPFFR2 interferes with the opioid system in the brain and its activation contributes to anxiety-like behavior in mice (Lin et al., 2017). Intraventricular application of NPFF reduced cocaine-induced conditioned place preference in rats pointing toward an important role of NPFF signaling in cocaine reward processing (Kotlinska et al., 2008). The *KALRN* gene encodes several isoforms and Kalirin-7 was shown to be involved in the development and function of dendritic spines (Ma et al., 2008). After repeated cocaine exposure in a mouse model, expression levels of Kalirin-7 were elevated in the NAc contributing to the formation of new dendritic spines. Kalirin-7 knock-out mice displayed increased cocaine-induced

locomotor sensitization and a decreased conditioned place preference (Kiraly et al., 2013).

Looking at multiple brain regions allows a more complete understanding of the role of DNAm alterations in CUD. In line with results from previous studies on epigenome-wide alterations of CUD in the CN (Vaillancourt, Yang, et al., 2021) and in the NAc (Vaillancourt, Chen, et al., 2021), we observed more hypermethylation than hypomethylation at the DMR level in BA9. In our systematic comparison of the DMRs that were consistently associated with CUD in both the CN and NAc, none were detected in our analysis. We detected one DMR annotated to the *ALDH3B1* that was conserved between BA9 and the CN. Hypermethylation of this DMR was also consistent between studies. It has to be noted that the EPIC array covers only a finite number of CpG sites which may limit the comparison with results from reduced representation bisulfite sequencing. In addition, we also compared our results with DNAm signatures of CUD in whole blood. While some nominally significant overlap was observed, there was little evidence for a systematic relationship between the CUD-associated CpG sites in BA9 and whole blood, as the overlap was not significant after multiple testing correction. DNAm is highly tissue-specific, which could partly explain the missing overlap. At the same time, the samples were phenotypically different in age and substance use.

In our analysis in BA9, no enrichment of hypo- and hypermethylated CUD-associated CpG sites among GO terms remained statistically significant after multiple testing correction. Further, pathway enrichment was driven by relatively few genes leading to the emergence of multiple biologically related pathways among the results. The functional relevance of CUD-associated differential methylation should thus be evaluated together with the enriched biological processes and hub genes derived from WGCNA co-methylation modules. Three of the four CUD-associated modules were enriched for neuron-specific processes such as synaptic signaling, ion channel activity, and neurotransmitter transport. Next to the association with CUD, modules brown and blue displayed a strong correlation with neuronal fraction, which might also contribute to the enrichment of neuronal processes within these co- methylation modules. PPI networks based on module hub genes revealed *CACNA1C* and *NR3C1* as highly connected genes for which a previous relation to cocaine use is known. *CACNA1C* is a well-described risk gene for psychiatric disorders such as bipolar disorder, MDD, and schizophrenia (Dedic et al., 2018). By encoding a subunit of the voltage-gated Cav1.2 calcium channel, it depicts a key regulator of neuronal excitability. Further, the expression of cocaine-induced locomotor sensitization was shown to be dependent on Cav1.2 signaling in mice (Giordano et al., 2010). The *NR3C1* gene encodes a glucocorticoid receptor for which downregulation in blood was observed in chronic cocaine users (Schote et al., 2019) and its DNAm levels were shown to

predict substance use phenotypes in adolescents (Raffetti et al., 2021). Further, brain-specific conditional knock-out of *NR3C1* flattened the dose-response curve for cocaine self-administration and reduced behavioral sensitization in an intravenous cocaine self-administration model in mice (Deroche-Gamonet et al., 2003). Thus, *CACNA1C* and *NR3C1* depict important regulators of the reinforcing effects of cocaine.

We further identified *JUN* as the hub node with the strongest connectivity in module steelblue. By forming heterodimers with FOS family proteins, JUN is involved in the formation of the AP- 1 transcription factor complex, a key regulator of neuroplasticity in addiction (McClung & Nestler, 2008). In a mouse model, decreased cocaine-induced conditioned place preference was found during expression of a dominant-negative mutant of c-Jun in the brain (Peakman et al., 2003). Epigenetic dysregulation of the JUN interaction network might therefore contribute to neurocircuit alterations in CUD.

In our analysis of epigenetic age in the brain, we observed a trend toward a more positive ageAccel in CUD cases compared to individuals without CUD that was more pronounced with Levine's clock compared to Horvath's clock. This is in line with results from previous studies that report on positive ageAccel, particularly for Levine's epigenetic clock, in the PFC of individuals with opioid intoxication, AUD, and OUD (Cabrera-Mendoza et al., 2023; Shu et al., 2021). As cocaine intake is associated with neuroinflammatory processes in the brain (López-Pedrajas et al., 2015; Sil et al., 2019), ageAccel in CUD could be a consequence of an inflammatory response that is particularly reflected by Levine's (Levine et al., 2018) epigenetic clock. However, as comorbid diseases are frequently observed in CUD, an increase in ageAccel could at least in part be driven by systemic diseases that are themselves associated with an accelerated epigenetic age. Due to the lack of information on somatic comorbidities, we were not able to address this point in the present study. Further, for the interpretation of epigenetic age in the brain, it must be noted that epigenetic clocks are at least to some extent tissue-specific and might not be applicable to every tissue. We only assessed epigenetic clocks that were either trained on non-blood tissues (Horvath) or at least showed a high correlation of findings across tissues (Levine). The negative values of Levine's PhenoAge we observed in some of the brain samples might represent a post-mortem brain tissue artifact, as Levine's clock was initially trained on blood methylation data. Similarly, in a recent study in post-mortem human brain tissue, negative PhenoAge estimates were observed for Levine's clock (Cabrera-Mendoza et al., 2023). Within-subject comparisons using blood and brain tissue from the same person should be performed to investigate such potential tissue artifacts. Further, analyses with larger sample sizes, careful adjustment for comorbidities, and appropriate estimation methods are required to evaluate if accelerated epigenetic aging in the brain is a true phenomenon in CUD.

Several limitations apply to our study on DNAm signatures of CUD in post-mortem human brain tissue. First, our sample size of $N = 42$ is comparatively low for an epigenome-wide analysis and is associated with limitations in statistical power. Another limitation is the frequent comorbidity of CUD with other SUDs and mood disorders, which is well reflected in our sample. Drug intoxication at death and medication prior to death could have affected methylation levels and therefore might confound the results for CUD. Next to the CUD phenotype, assessment of used substances and comorbidities including their treatment is crucial to adjust for confounding effects on DNAm signatures. Due to the small sample size, we were not able to adjust for drug intoxication at death and medication prior to death. Thus, we cannot rule out a potential confounding of the identified CUD-associated methylation signatures by epigenetic effects of other substances. Third, the case-control design of the EWAS does not allow the assessment of temporal changes in DNAm levels. Thus, within the present study, we are not able to separate differential methylation acquired during the disease course of CUD from an epigenetic predisposition to CUD.

In the present study, we identified CUD-associated DNAm signatures in the BA9 subregion of the human PFC. We detected CUD-associated differential methylation within several transcription factor genes, an enrichment within genomic features of transcription regulation, and an association with gene regulatory networks involved in neurotransmission and neuroplasticity. Specifically, epigenetic alterations related to calcium and cAMP signaling as well as the *CACNA1C* and *JUN* PPI networks might contribute to the development of the altered neurocircuits observed in the CUD brain.
To follow up on these hypotheses, further studies are required. These should be based on larger sample sizes with detailed phenotyping to minimize the confounding of results in association studies. Further, integration of epigenome-wide methylation data with other omics data such as transcriptomics or proteomics is necessary to prioritize and substantiate findings from the EWAS. Such multi-omics studies depict promising approaches for deeper profiling of altered biological processes in SUDs as recently shown for AUD (Zillich, Poisel, et al., 2022) and OUD (Mendez et al., 2021). To address the temporal dynamics of molecular changes in the disease course of CUD, a longitudinal study assessing different stages of CUD within the same individuals should be performed. While a longitudinal design is nearly impossible using post-mortem brain tissue, it is a feasible approach for blood samples. The application of such innovative methods could pave the way for the development of precision medicine approaches thereby addressing the urgent need for novel therapeutic options in CUD.

2.6 Data availability statement

Methylation data presented in this study have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number: EGAS00001006826 (https://ega-archive.org/studies/EGAS00001006826).

2.7 Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee II of the Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany under the register number: 2021-681. Written informed consent was not provided because post-mortem brain tissue was obtained from individuals that died by suicide or sudden death. Following the established ethical standards of the DBCBB, written informed consent was obtained from the next-of-kin for each subject.

2.8 Author contributions

SW, RS, AH, and MR: conceptualization and funding acquisition. EP, FS, JCF, and LZ: data curation. EP and LZ: formal analysis and writing—original draft. EP, LZ, JCF, MF, AH, RS, and SW: investigation. LZ, EP, FS, and JF: methodology. SW, MN, AH, and RS: project administration. GT, NM, AH, and MN: resources. SW, JF, MR, and RS: supervision. All authors contributed to the article writing—review and editing and approved the submitted version.

2.9 Funding

Funding was provided by the German Federal Ministry of Education and Research (BMBF) through the e:Med research program "A systems-medicine approach toward distinct and shared resilience and pathological mechanisms of substance use disorders" (01ZX01909 to RS, MR, SW, and AH). Further, by "Toward Targeted Oxytocin Treatment in Alcohol Addiction (Target-OXY)" (031L0190A to JCF), the ERA-NET program: Psi- Alc (FKZ: 01EW1908), the Hetzler Foundation for Addiction Research (to AH), and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 402170461–TRR 265 (Heinz et al., 2020) to RS, AH, and MR.

2.10 Acknowledgements

We thank Claudia Schäfer-Arnold and Elisabeth Röbel for technical assistance.

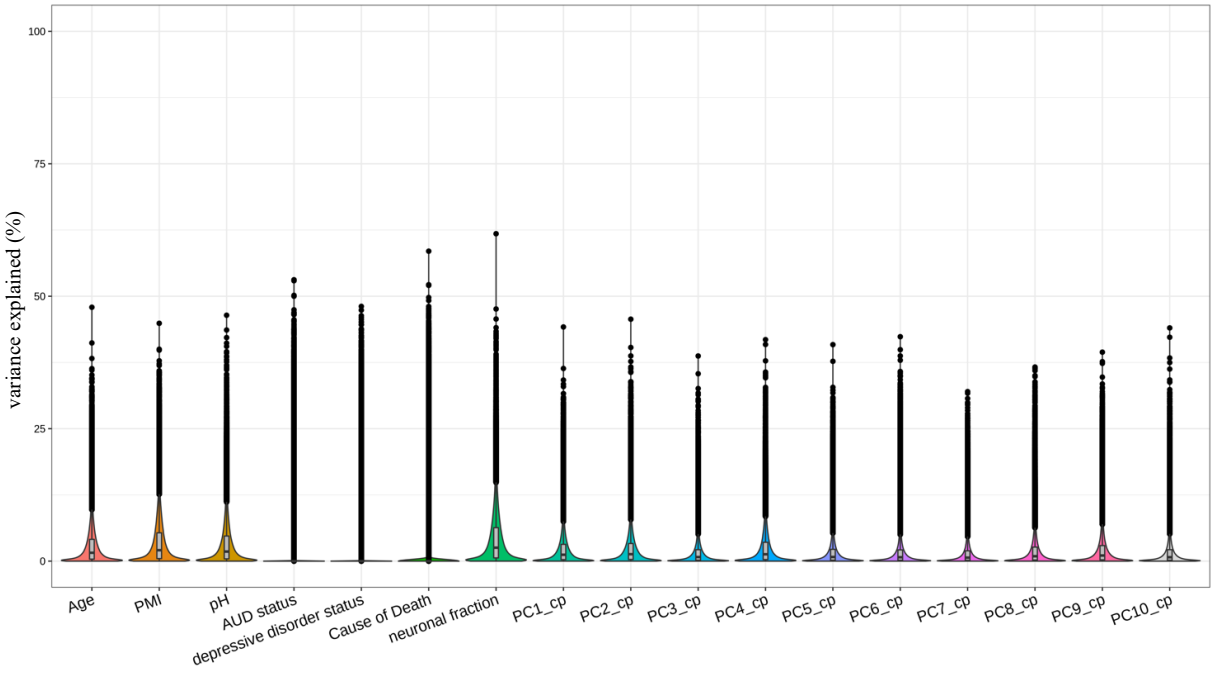
2.11 Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

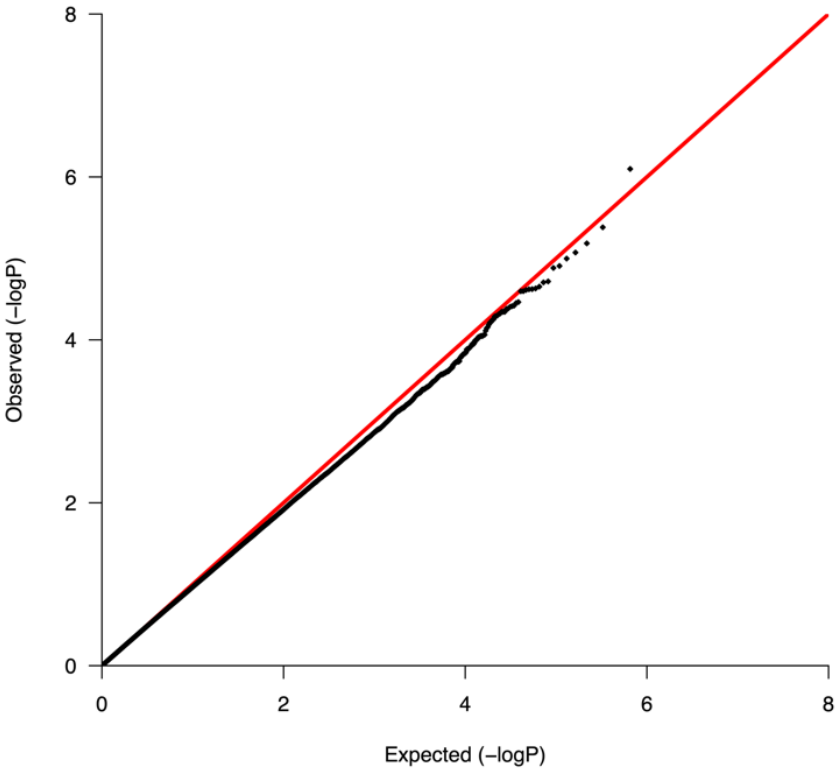
2.12 Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2023.1075250/full#supplementary-material

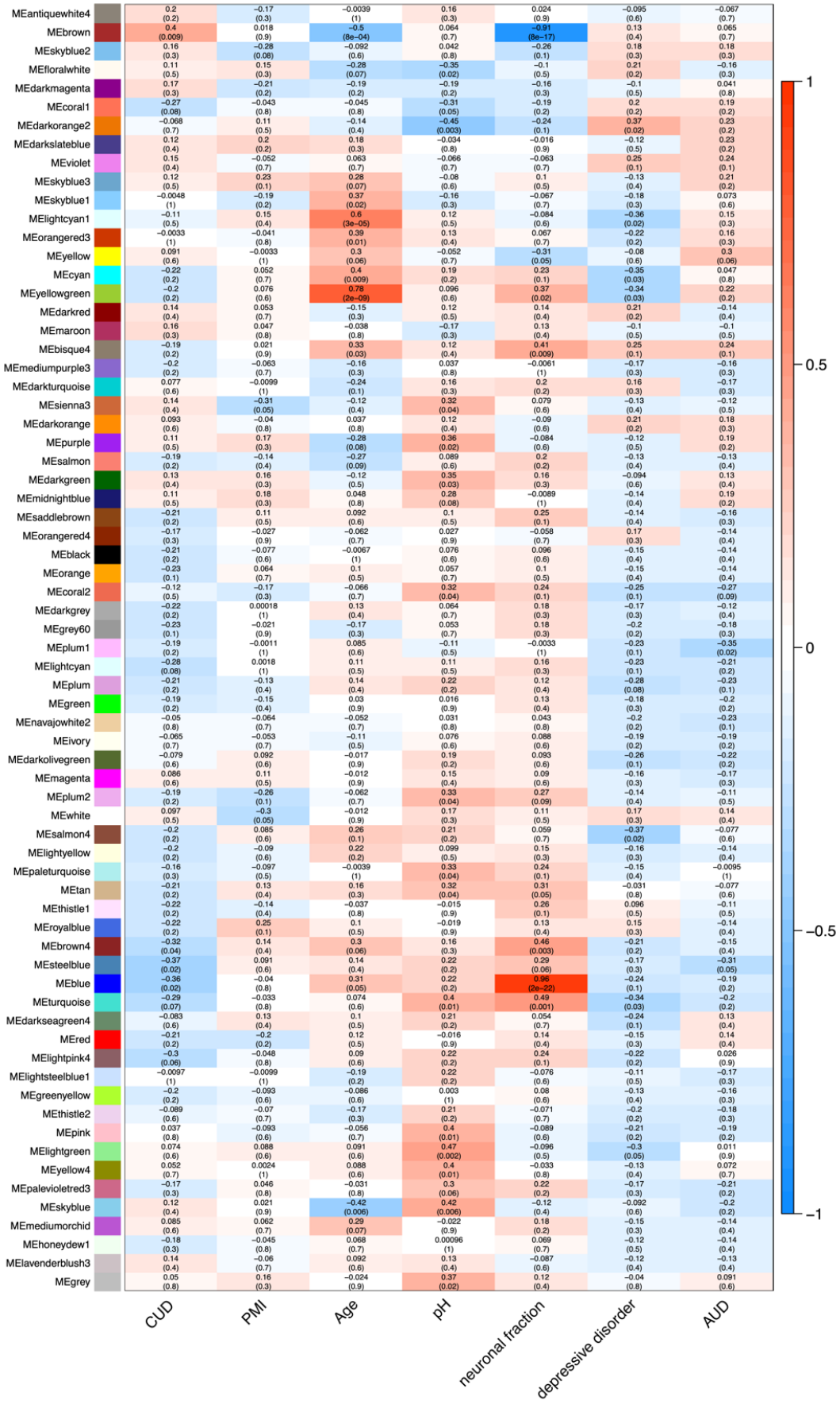
2.12.1 Supplementary Figures



Supplementary Figure 2.1 - Results of the variance partition analysis confirming the selected covariates for the linear model in the epigenome-wide association study.

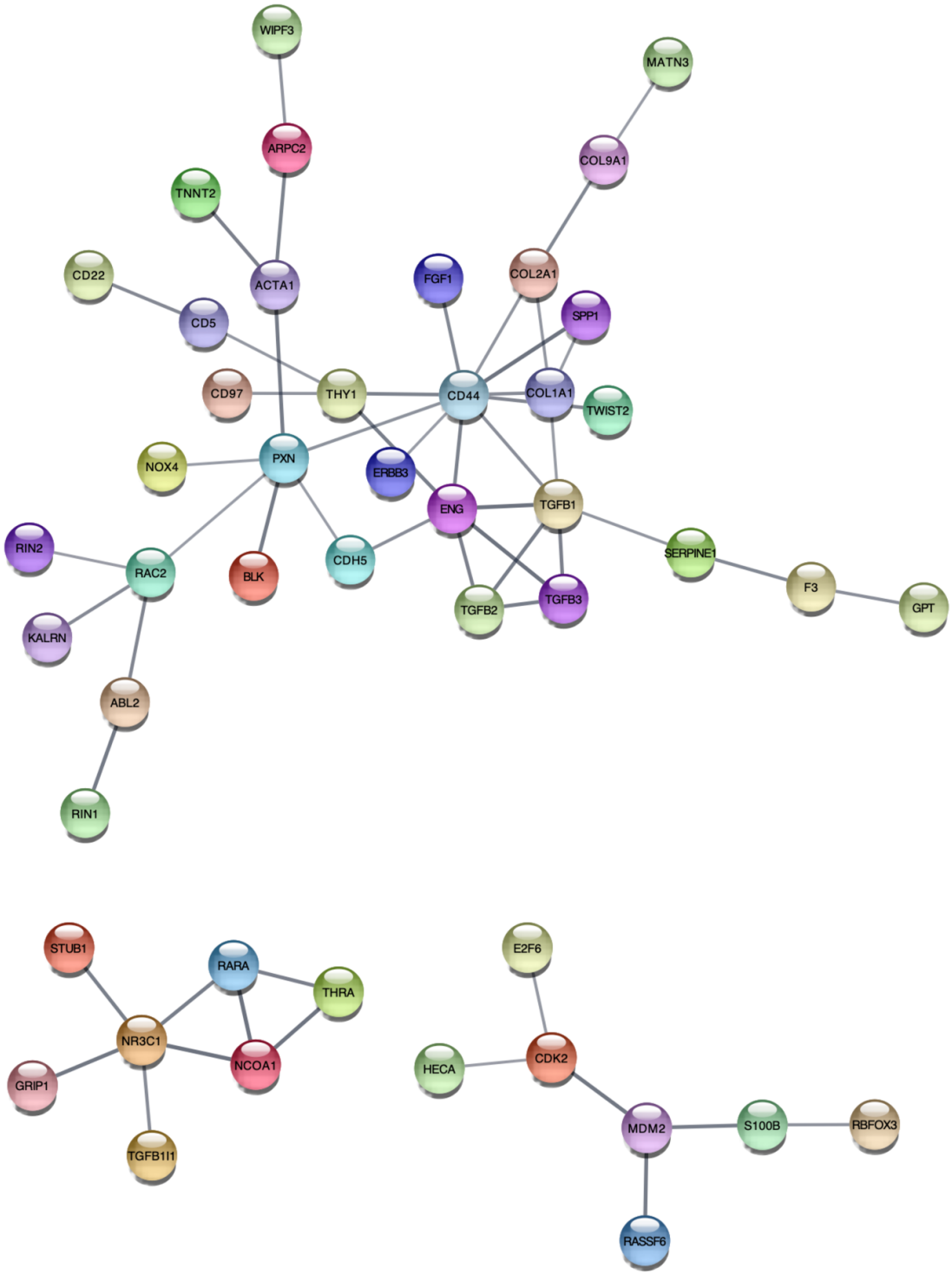


Supplementary Figure 2.2 - QQ-Plot (λ = 0.969) of the EWAS of cocaine use disorder in human postmortem brain tissue of Brodmann Area 9 (N = 42).

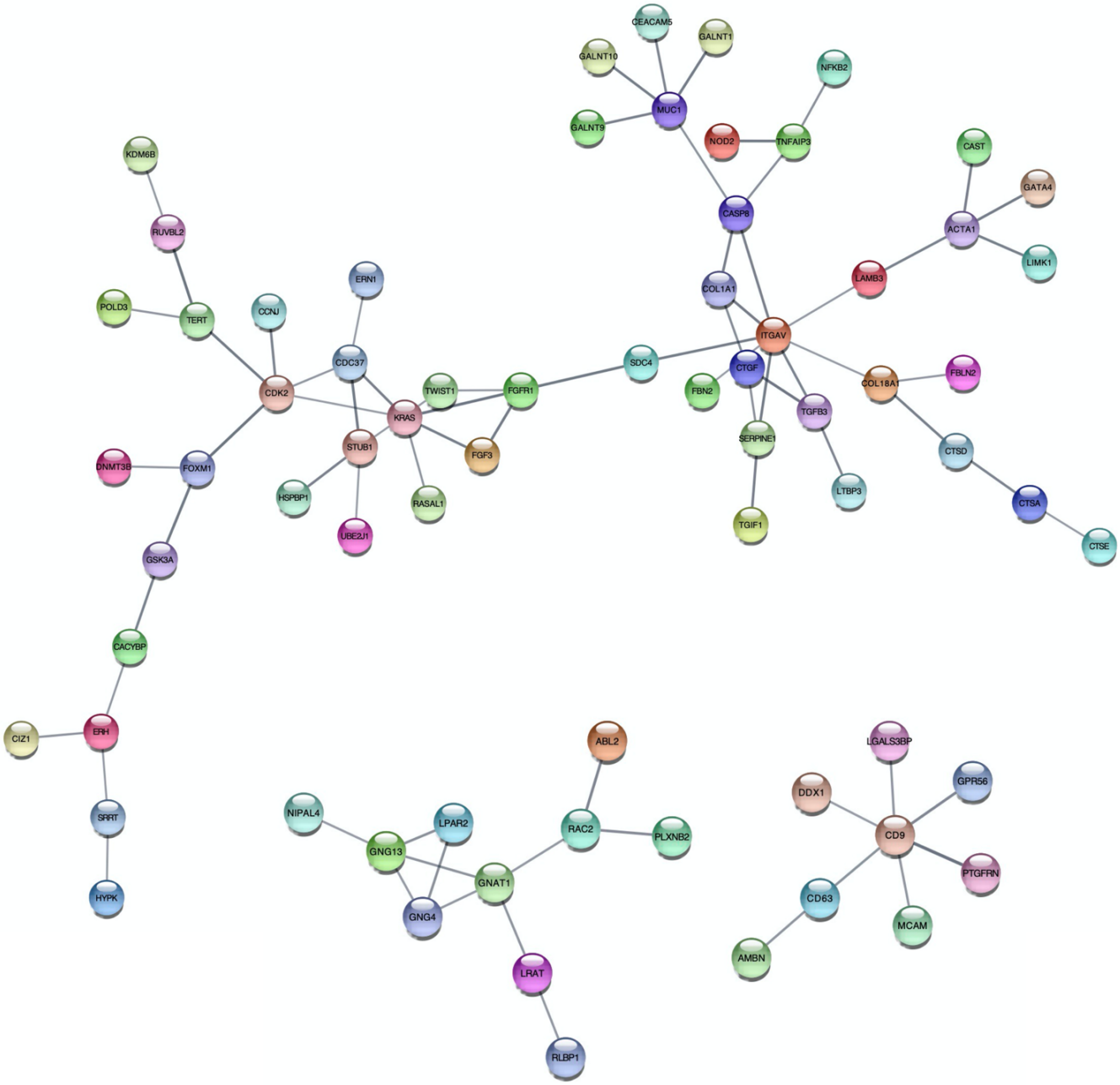


Supplementary Figure 2.3 - Module-trait correlation plot displaying all co-methylation modules resulting from WGCNA. CUD: cocaine use disorder, PMI: postmortem interval, neuronal fraction: estimated neuronal cell type proportion, AUD: alcohol use disorder.

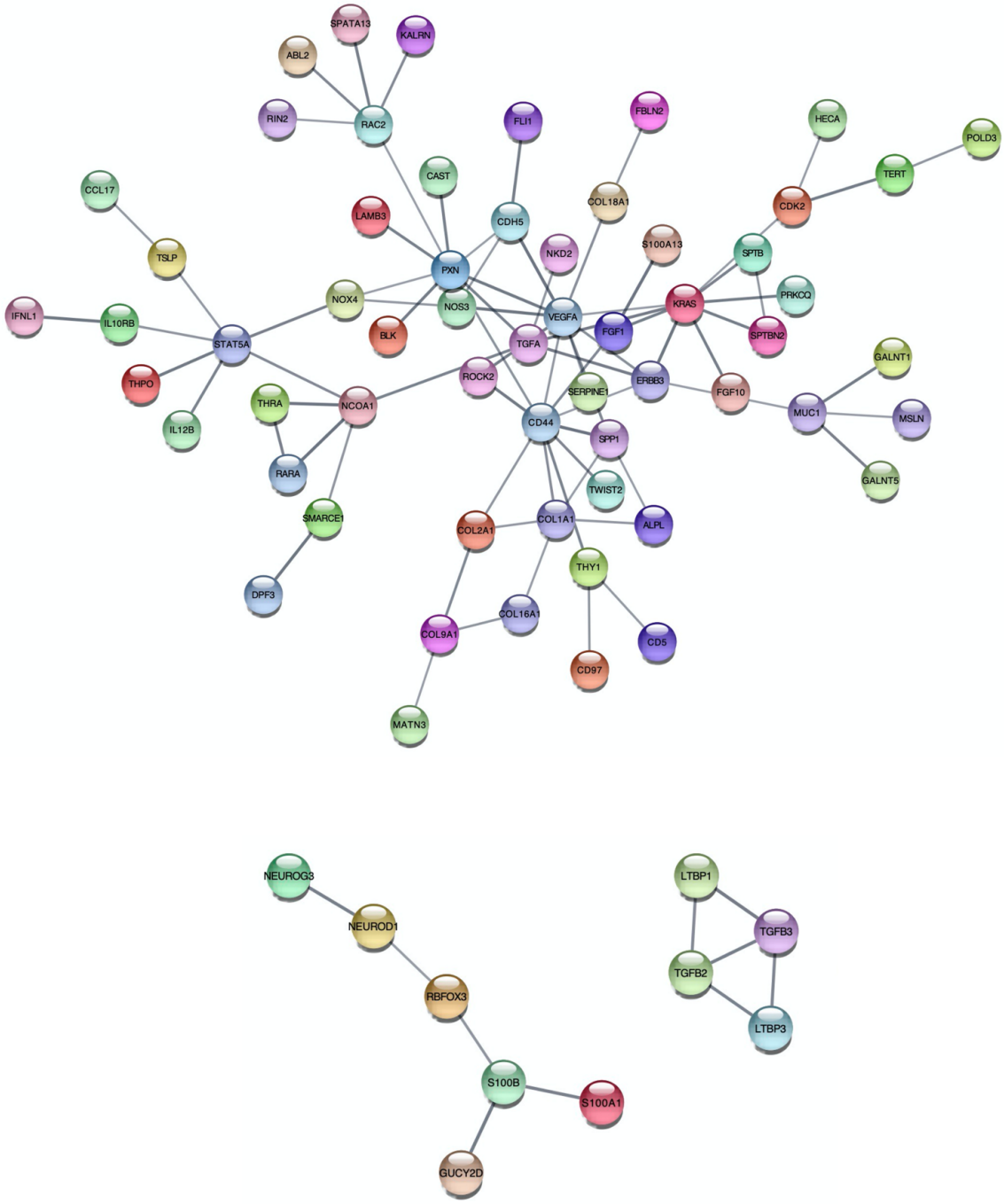
A module brown



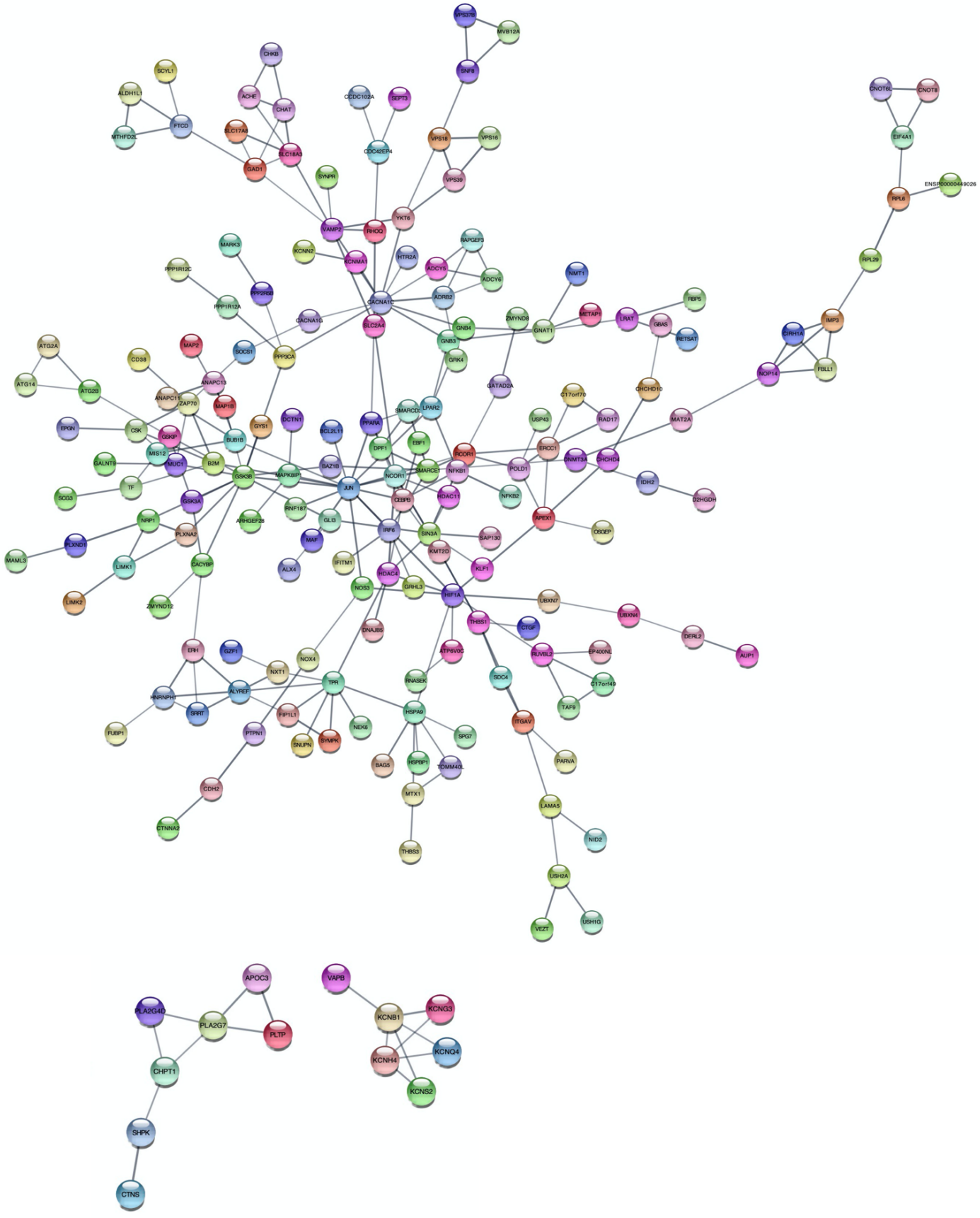
B module brown4



C module blue



D module steelblue



(Figure legend on next page)

Supplementary Figure 2.4 - Protein-protein-interaction networks derived from CUD-associated co-methylation module hub genes. The top three networks ranked by size and the connectivity of nodes are shown for WGCNA-derived co-methylation modules **A** brown, **B** brown4, **C** blue, and **D** steelblue. Network plots were generated using the *Search Tool for the Retrieval of Interacting Genes/Proteins* (STRING, v.11.5) with an interaction score threshold of 0.7 (high confidence interactions).

2.12.2 Supplementary Tables

Due to genome-wide table sizes, for full supplementary tables, please refer to the online version of the supplementary information:

https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2023.1075250/full#supplementary-material

Supplementary Table 2.1. Detailed information on cause of death, toxicology at death and medication during the last 3 months prior to death. GHB: gamma-hydroxybutyric acid, p-value: derived from CUD/ no CUD.

Supplementary Table 2.2. Summary statistics of the EWAS for the top 1000 CUD-associated CpG sites. FDR: FDR significance after multiple testing correction, meanM: mean M value of methylation for the respective CpG site.

Supplementary Table 2.3a. EWAS summary statistics from the analysis in BA9 for the available CUD-associated CpG sites identified in whole blood (Camilo et al., 2019, Braz J Psychiatry, Supplementary Material). FDR: FDR significance after multiple testing correction, meanM: mean M value of methylation for the respective CpG site.

Supplementary Table 2.3b. Associations with CUD for CpG sites in 500 base pairs proximity to DMRs identified in both the caudate nucleus and nucleus accumbens (Vaillancourt et al., 2021, iScience, and Vaillancourt et al. 2021, Mol Psychiatry). meanM: mean M value of methylation for the respective CpG site.

Supplementary Table 2.4a. Results of the GO enrichment analysis based on CpG sites associated ($p<0.001$) with CUD. Overrepresentation of CpG sites with positive (hypermethylation) and negative effect sizes (hypomethylation) among GO terms was tested separately. The top 20 pathways are displayed.

Supplementary Table 2.4b. Results of the GO enrichment analysis for CpG sites located in the N=20 CUD-associated DMRs in BA9 (goregion function in missMethyl). The top 20 pathways showing strongest overrepresentation of CpGs annotated to genes are displayed.

Supplementary Table 2.5. Results of the GO enrichment analysis in CUD-associated co-methylation modules from WGCNA. The 10 GO terms with strongest overrepresentation are shown.

Supplementary Table 2.6. Module hub genes from CUD-associated WGCNA co-methylation modules (top 0.5% of genes ranked by product of gene significance for CUD status and module membership).

Supplementary Table 2.7a. Module hub genes from WGCNA co-methylation module brown - node table with stringdb identifiers resulting from the final STRING network.

Supplementary Table 2.7b. Module hub genes from WGCNA co-methylation module brown4 - node table with stringdb identifiers resulting from the final STRING network.

Supplementary Table 2.7c. Module hub genes from WGCNA co-methylation module blue - node table with stringdb identifiers resulting from the final STRING network.

Supplementary Table 2.7d. Module hub genes from WGCNA co-methylation module steelblue - node table with stringdb identifiers resulting from the final STRING network.

Supplementary Table 2.8. Top 10 highly connected nodes (hub nodes) of STRING protein-protein interaction networks derived from hub genes of the CUD-associated WGCNA modules.

Supplementary Table 2.9. Results from the association analysis of CUD with epigenetic aging using linear models.

3 STUDY 2 - MULTI-OMICS PROFILING OF DNA METHYLATION AND GENE EXPRESSION ALTERATIONS IN HUMAN COCAINE USE DISORDER[2]

3.1 Abstract

Structural and functional changes of the brain are assumed to contribute to excessive cocaine intake, craving, and relapse in cocaine use disorder (CUD). Epigenetic and transcriptional changes were hypothesized as a molecular basis for CUD-associated brain alterations. Here we performed a multi-omics study of CUD by integrating epigenome-wide methylomic ($N = 42$) and transcriptomic ($N = 25$) data from the same individuals using postmortem brain tissue of Brodmann Area 9 (BA9). Of the $N = 1\ 057$ differentially expressed genes ($p < 0.05$), one gene, *ZFAND2A*, was significantly upregulated in CUD at transcriptome-wide significance ($q < 0.05$). Differential alternative splicing (AS) analysis revealed $N = 98$ alternatively spliced transcripts enriched in axon and dendrite extension pathways. Strong convergent overlap in CUD-associated expression deregulation was found between our BA9 cohort and independent replication datasets. Epigenomic, transcriptomic, and AS changes in BA9 converged at two genes, *ZBTB4* and *INPP5E*. In pathway analyses, synaptic signaling, neuron morphogenesis, and fatty acid metabolism emerged as the most prominently deregulated biological processes. Drug repositioning analysis revealed glucocorticoid receptor targeting drugs as most potent in reversing the CUD expression profile. Our study highlights the value of multi-omics approaches for an in-depth molecular characterization and provides insights into the relationship between CUD-associated epigenomic and transcriptomic signatures in the human prefrontal cortex.

3.2 Introduction

Cocaine use disorder (CUD) is a globally prevalent substance use disorder (SUD) with around 4.2 million people worldwide being diagnosed with CUD (Degenhardt et al., 2018). Individuals suffering from CUD present with compulsive cocaine use patterns, strong cocaine craving, and high rates of relapse even after prolonged time of abstinence (American Psychiatric Association, 2013). Currently, there is no FDA-approved pharmacotherapy for CUD and treatment is mainly focused on symptom reduction (Kampman, 2019). Neurobiological alterations in the brain are

[2] Publication: Zillich, E., Belschner, H., Avetyan, D., Andrade-Brito, D., Martínez-Magaña, J. J., Frank, J., Mechawar, N., Turecki, G., Cabana-Domínguez, J., Fernàndez-Castillo, N., Cormand, B., Montalvo-Ortiz, J. L., Nöthen, M. M., Hansson, A. C., Rietschel, M., Spanagel, R., Witt, S. H., & Zillich, L. (2024). Multi-omics profiling of DNA methylation and gene expression alterations in human cocaine use disorder. *Translational Psychiatry*, *14*(1), 428. https://doi.org/10.1038/s41398-024-03139-9

assumed to contribute to the observed clinical symptoms in CUD (Nestler, 2005). This is supported by neuroimaging studies that have shown profound structural and functional alterations in the brain in individuals with CUD (Ersche et al., 2011; Goldstein & Volkow, 2002). In addition to striatal brain regions involved in reward processing (Cox & Witten, 2019), frontal cortical areas that are neuroanatomically connected with limbic structures, are implicated in addiction due to their importance for inhibitory control (Ersche et al., 2011; Goldstein & Volkow, 2002; Koob & Volkow, 2016).

Dynamic changes in epigenetics and gene expression were hypothesized as a molecular basis of CUD-associated brain changes (Fernàndez-Castillo et al., 2022; Robison & Nestler, 2011). So far, the majority of studies investigating brain tissue focused on rodent models of cocaine addiction, identifying specific genomic loci to be differentially methylated in brain regions such as the prefrontal cortex (PFC) (Baker-Andresen et al., 2015) and nucleus accumbens (NAc) (Anier et al., 2010). The prefrontal cortex with subregions, such as Brodmann Area 9 (BA9) is of particular interest in studying SUDs as it's involved in executive control and, hence, plays a major role in the preoccupation/anticipation stage in the neurocircuitry of addiction (Koob & Volkow, 2016). Gene expression levels are tightly regulated by epigenetic mechanisms and DNA methylation (DNAm) changes especially in gene promoter regions were shown to alter transcript abundance (Gibney & Nolan, 2010). In line with this, differential gene expression in rodent models of cocaine addiction was reported in multiple brain regions where transcription factors of the immediate early gene (IEG) family such as *Egr1*, *Nr4a1*, and *Fos* were found to be differentially expressed (Campbell et al., 2021; Li et al., 2017; Savell et al., 2020; Teague & Nestler, 2022). At the transcriptome-wide scale, differentially expressed genes were consistently enriched in biological processes related to neurotransmission and ion channel activity, but also metabolic alterations related to lipid metabolism and ATP homeostasis were found (Campbell et al., 2021).

Few studies have been performed investigating genome-wide DNAm or transcriptomic changes in CUD in human postmortem brain tissue. Two epigenome-wide studies using reduced representation bisulfite sequencing (RRBS) in a cohort of $N = 25$ individuals with CUD and $N = 25$ control individuals identified $N = 145$ and $N = 173$ CUD-associated differentially methylated regions (DMRs) in the NAc (Vaillancourt, Chen, et al., 2021) and in the caudate nucleus (CN) (Vaillancourt, Yang, et al., 2021), respectively. Investigating the same brain regions in a different cohort ($N = 25$ CUD cases, $N = 20$ controls), another study characterized transcriptome-wide gene expression changes and reported on the upregulation of synaptic transmembrane transporter genes while immune processes were downregulated (Mews et al., 2023). The largest study in the human PFC investigating CUD-associated transcriptomic changes ($N = 19$ CUD, $N = 17$

controls) identified $N = 883$ nominally significant ($p < 0.05$) differentially expressed genes (DEGs) in neuronal nuclei from the Brodmann Area 46 subregion (Ribeiro et al., 2017). CUD-associated co-expression networks were enriched for GTPase signaling and neurotransmitter secretion. Regarding epigenomic alterations in the PFC, we were previously able to identify 20 CUD-associated DMRs in Brodmann Area 9, a subregion of the PFC, and further detected that co-methylation networks in CUD were enriched for synaptic signaling processes (Poisel et al., 2023). Although epigenetics represents an important regulatory mechanism for transcription, the co-regulation of DNAm and gene expression in the same brain samples has not yet been investigated in CUD, limiting the comparability of results between epigenetic and gene expression studies.

In addition to epigenetics and transcription, alterations of alternative splicing might contribute to the neurobiological changes in the CUD brain, as shown in other SUDs. Previous studies using postmortem human brain tissue from individuals with alcohol use disorder (AUD) (Huggett et al., 2023; Li et al., 2023; Van Booven et al., 2021) and opioid use disorder (OUD) (Huggett et al., 2022) detected differential alternative splicing in transcripts of genes implicated in neuropsychiatric disorders, such as *BIN1*, *FLOT1*, and *ELOVL7* suggesting RNA splicing alterations to be a further molecular mechanism in the neurobiology of SUDs. While a recent study using a cocaine self-administration model in mice showed widespread changes in alternative splicing in multiple brain regions (Xu et al., 2021), no systematic evaluation of splicing alterations in human CUD was performed so far.

In the present study, we aimed to characterize the molecular underpinings of CUD in the human prefrontal cortex by applying a multi-omics analysis approach. We investigated differentially expressed genes in postmortem brain tissue from deceased CUD cases compared to well-matched controls and integrated them with the results of our epigenome-wide DNAm analysis from the same individuals of the BA9 subregion of the human PFC (Poisel et al., 2023). Further, we characterized differential alternative splicing in BA9. We then performed replication analysis of CUD-associated DEGs in two other independent RNA-seq datasets of human dlPFC. Gene expression data, including alternative splicing results, and DNA methylation data were then integrated and put into a biological context. Finally, we addressed the urgent need for novel therapeutic approaches, by performing a drug repositioning analysis based on the CUD-associated transcriptional profile in BA9.

Collectively, our multi-omics study design represents an integrated analysis of DNAm and gene expression data together with alternative transcript splicing that

highlights the role of synaptic and metabolic alterations in CUD and the glucocorticoid receptor as a pharmacological candidate target.

3.3 Methods

3.3.1 Postmortem human brain tissue

The sample of human postmortem brain tissue of BA9 was obtained from the Douglas Bell Canada Brain Bank (DBCBB). Inclusion criteria were age >18 and a diagnosis of cocaine dependence based on DSM-IV. Throughout this study, we will nevertheless use the more recent terminology from DSM-5 i.e., cocaine use disorder. Individuals were excluded from the study if they were diagnosed with severe neurodevelopmental or psychiatric disorders other than depressive disorders or had received additional diagnoses of substance use disorders other than alcohol use disorder. All included subjects were male and of European American descent. Demographic information for the cohort of $N = 42$ BA9 tissue donors is described in Table S3.1 and for the subset of $N = 25$ individuals with RNA-seq data in Table 3.1.

3.3.2 DNA methylation data generation

DNA extraction was performed as described in (Poisel et al., 2023). In brief, DNA was extracted from the full set of $N = 42$ BA9 samples using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). The epigenome-wide DNAm profile was determined using the Illumina MethylationEPIC BeadChip v1 (850k) (Illumina, San Diego, CA, USA). During sample processing and analysis of DNAm levels, randomization was applied based on CUD status and known comorbidities such as AUD and depressive disorders.

3.3.3 Generation of gene expression data

Using the miRNeasy Tissue/Cells Advanced Micro Kit (Qiagen, Hilden, Germany), total RNA was extracted from the $N = 42$ BA9 samples using ~5 mg of frozen tissue per sample. The RNA integrity number (RIN) was measured using a TapeStation 4200 (Agilent, Santa Clara, CA, USA) resulting in a total of $N = 25$ samples remaining for RNA sequencing (RIN > 5.5). Following ribosomal RNA (rRNA) depletion, libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) followed by sequencing with an average of 60 million read pairs (2x100bp) per sample. RNA sequencing was performed using an Illumina NovaSeq 6000 device.

3.3.4 Statistical analyses

All statistical analyses in the R programming environment were performed using R version 4.2.1. If not otherwise stated, adjustment for multiple testing was

performed using the Benjamini-Hochberg (FDR) procedure (Benjamini & Hochberg, 1995). An analysis workflow for the multi-omics study of DNA methylation and gene expression in CUD is shown in Supplementary Fig. S3.8.

3.3.5 DNA methylation analysis

Methylation data was analyzed as part of the Poisel et al. (2023) study where a detailed description of the analysis pipeline can be found in the methods section. In brief, DNA methylation levels were preprocessed using an in-house quality control (QC) pipeline based on CPACOR (Lehne et al., 2015). The neuronal cell fraction was estimated based on the Houseman algorithm (Houseman et al., 2012) using a dlPFC reference dataset (Jaffe & Kaminsky, 2017). Quantile-normalized beta values were derived from raw-intensities, followed by logit-transformation to M values of methylation. An epigenome-wide association study (EWAS) was performed using a linear regression model while adjusting for covariates that have a known effect on DNA methylation such as age, postmortem interval (PMI), pH of the brain tissue, neuronal cell fraction, comorbid depressive and/or alcohol use disorder, and technical factors. Downstream analyses based on the results of the EWAS included the identification of differentially methylated regions (DMRs), a gene ontology enrichment analysis using CUD-associated CpG sites ($p_{assoc}<0.001$), and a network analysis in WGCNA to evaluate CUD-associated co-methylation modules.

3.3.6 Gene expression analysis

Sequencing quality metrics were inspected using FastQC v.0.12.1 confirming all 25 fastq files to be used in further analysis. Reads were mapped to the GRCh38 genome primary assembly using STAR v.2.7.10b (Dobin et al., 2013). Quantification of features was performed using the featureCounts implementation in the R package Rsubread v.2.12.3 (Liao et al., 2019) with the genome annotation gtf-file v.43 from GENCODE (https://www.gencodegenes.org). The raw count matrix was imported in DESeq2 v.1.38.3 (Love et al., 2014) and differential expression (DE) testing was performed while adjusting for the covariates age, PMI, brain pH and RIN resulting in the following DESeq2 design formula: mRNA ~ CUD + age + PMI + pH + RIN. Fold-change cut-offs for DEGs were an absolute log2 fold change of larger than 0.07, corresponding to a 5% change in transcript abundance. Statistical significance cut-offs were $p < 0.05$ for nominal significance and $q < 0.05$ for a 5% FDR-adjusted significant association with CUD. All covariates included in the DESeq2 model are known to influence the gene expression profile and were confirmed in a variance partition analysis in our dataset using the R package variancePartition v.1.28.7 (Supplementary Fig. S3.1a). Variables characterized by median variance explanation (var.exp) larger than 0.01 across all transcripts were included as covariates into the statistical model. As comorbid MDD and AUD explained only minimal variance in

the expression data (median var.exp=0) and only 25 of the 42 samples were available in the expression analysis, MDD and AUD were not included as covariates in the statistical model. A sensitivity analysis was performed including MDD and AUD as covariates (Supplementary Fig. S3.1b) confirming a strong correlation between the log2 fold-changes of the nominally significant results.

3.3.7 Cell type deconvolution analysis

Based on reference signatures of gene expression derived from single-cell studies, the distribution of cell types in bulk expression data can be inferred using cell type deconvolution algorithms such as CIBERSORT (Newman et al., 2015). We used a curated set of cell type-specific marker genes of the human prefrontal cortex based on a study from Yu and He (Yu & He, 2017) where a gene was required to have a 10-fold stronger expression in a specific cell type compared to all other cell types to be considered a marker gene. DESeq2-normalized counts of the BA9 expression dataset were used and cell type deconvolution was performed using the CIBERSORT R script v1.04. To test for significant differences in cell type distribution in samples from individuals with and without CUD, we performed a Bayesian estimation of the difference in means and evaluated the 95% high-density interval. The Bayesian testing was based on BEST (Kruschke, 2013) as implemented in the R package BayesianFirstAid v.0.1. Further, an overlap analysis of DEGs in cell type markers was performed in GeneOverlap v.1.34.0 (Shen L, 2023) using the 10-fold marker gene list from (Yu & He, 2017) in a Fisher test.

3.3.8 Functional enrichment analysis

To characterize altered biological functions related to the observed gene expression differences, we performed a gene set enrichment analysis (GSEA) for Gene Ontology (GO) terms using the gseGO function from the R package clusterProfiler v.4.6.2 (Yu et al., 2012). The DESeq2 Wald statistic defined as the log2FC divided by its standard error was used for ranking of the results. A significance threshold of $q < 0.05$ (5% FDR) was considered statistically significant. Results of the GSEA were visualized using the emapplot function in enrichplot v.1.18.3.

3.3.9 Weighted gene co-expression network analysis

To identify CUD-associated co-expression patterns, we constructed co-expression modules using network analysis in weighted gene co-expression network analysis (WGCNA) R package v.1.72.1 (Langfelder & Horvath, 2008) and related them to CUD and other phenotypic variables available in the DBCBB cohort. Using the input matrix of normalized and variance stabilization transformed (vst) gene counts from DESeq2, a soft power threshold of 9 was estimated to achieve the criterion of scale free topology ($R^2 > 0.85$). For the construction of networks, we

used the parameters minModuleSize = 10, mergeCutHeight = 0.25, and maxBlockSize=36 000. The Pearson correlation of the module eigengene derived from each of the resulting $n = 27$ co-expression modules with the phenotypes of interest including CUD was calculated to identify significant associations of the modules with phenotypes (Fig. S3.3A). Downstream analyses of modules significantly associated with CUD included a GO enrichment analysis using the genes assigned to the modules using the full genome as the background. Next, module genes were ranked by the product of gene significance*module membership to identify hub genes. The top 10% of module hub genes were further investigated by constructing protein-protein interaction (PPI) networks. For this, Cytoscape v.3.9.1 (Shannon et al., 2003) with stringApp v.1.7.0 (Doncheva et al., 2019) was used. A detailed description of the PPI visualization settings in Cytoscape is found in (Poisel et al., 2023).

3.3.10 Replication analysis of differential expression results

Replication analysis of CUD-associated DEGs was performed in two independent datasets where RNA-seq data from postmortem human brain tissue of the prefrontal cortex from individuals with and without CUD was available. As the first replication dataset, BA9 bulk RNA-sequencing data from $N = 7$ individuals with CUD and $N = 14$ control individuals originating from the National PTSD Brain Bank (NPBB) (Friedman et al., 2017) was used. Phenotypic information for the BA9 replication cohort is shown in Table S3.17. RNA-seq data sequenced and pre-processed as described in (Girgenti et al., 2021) was analyzed for CUD-associated differential gene expression in DESeq2 using donor age, sex, PMI, and RIN as covariates. A comparative overview on variance explanation by covariates used in the analysis of discovery and replication cohorts is provided in Supplementary Fig. S3.4. The second replication cohort was based on a neuronal-specific RNA-sequencing dataset (GEO accession number: GSE99349) as described in (Ribeiro et al., 2017). In this study, neuronal nuclei were isolated from postmortem human brain tissue of the Brodmann Area 46 subregion of the dlPFC that is laterally adjacent to BA9. Here, bulk RNA-seq data was generated from $N = 19$ individuals with CUD and $N = 17$ without CUD from a male mixed ancestry cohort originating from the University of Miami Brain Bank (MBB). Raw sequencing data from the replication cohort was downloaded from GEO and processed using the same analysis pipeline as in the BA9 discovery sample: 1) mapping using STAR, 2) quantification using featureCounts, and 3) DE analysis in DESeq2. For the replication analysis in MBB data, we used the same statistical model as in the discovery analysis with differential expression testing for CUD while adjusting for donor age, RIN, pH, and PMI. To explore the results, we first performed an overlap analysis of nominally significant CUD-associated DEGs ($p < 0.05$) identified in the three datasets. Second, a targeted look-up of effect sizes

(log2FC) and association *p*-values was performed for overlapping DEGs across datasets and for the top findings from the BA9 discovery sample, *ZBTB4* and *INPP5E*. As an additional replication approach, we performed rank-rank hypergeometric overlap (RRHO) using the R package RRHO2 v.1.0 (Cahill et al., 2018) to evaluate convergent and divergent expression patterns at the transcriptome-wide scale between studies. RRHO scores were generated based on full differential expression statistics from discovery and replication datasets followed by the evaluation of overlapping signatures between studies using the hypergeometric testing procedure as implemented in RRHO2.

3.3.11 Signature-based drug repositioning analysis

With the top 150 upregulated and downregulated genes ranked by the DESeq2 test statistic from the differential expression analysis, the maximum input size in the Connectivity Map (CMap) query tool (https://clue.io/query, software version 1.1.1.43) was used to evaluate the connectivity of expression signatures (Table S3.8). CMap query uses the L1000 assay from the NIH LINCS project (https://lincsproject.org/) as a drug-gene expression relationship database. In L1000, expression changes for a representative set of 978 landmark transcripts are measured in response to treatment with a perturbagen such as a pharmaceutical drug (Subramanian et al., 2017). In addition to the connectivity scores for individual perturbagens, CMap also provides information on perturbagen classes and a GSEA output for pathways and drug targets. Normalized connectivity scores and FDR-adjusted *p*-values for perturbagens and GSEA results were obtained from the CMap query tool and visualized as waterfall plots in R using ggplot2 v.3.4.2.

3.3.12 Differential splicing analysis

Alternative splicing was evaluated using the annotation-free quantification approach of RNA splicing in LeafCutter v.0.2.9 (Li et al., 2018). The intron-centered approach of LeafCutter allows for differential splicing analyses in short-read sequencing datasets that have been difficult to access with previous alternative splicing analysis methods. First, raw sequencing data were aligned to the GRCh38 reference genome using STAR with an adapted 2-pass mapping procedure. For this, the first mapping step was performed using a regular gtf-file-derived genome index. The resulting splice junctions (SJ_out.tab-files) from the $N = 25$ samples were combined and filtered so that non-canonical junctions, junctions that were supported by less or equal than 2 uniquely mapping reads, annotated junctions already covered by the gtf-file, and duplicated junctions were removed. Using the filtered splice junction output, a modified genome index was derived using STAR in genomeGenerate mode. This extended genome index containing information on gene annotation and splice junctions was used in the

second mapping step resulting in the final bam-file output after mapping. Generation of junc-files, intron clustering, and differential intron excision analysis was performed as outlined by the authors of leafCutter (https://davidaknowles.github.io/leafcutter/) while including age, PMI, pH, and RIN as covariates into the Dirichlet-Multinomial generalized linear model. Default settings were used in the leafcutter_ds.R script i.e. maximum cluster size =10, minimum samples per intron = 5, minimum samples per grou*p* = 3, and a minimum coverage of 20 reads. The differential intron excision analysis results in an estimate for the change in the percent spliced in measure (ΔPSI) for each intron in a cluster and an FDR-adjusted *p*-value for the cluster in which the differential splicing events were detected. Differential splicing events in clusters with $|dPSI| > 0.025$ and an FDR-adjusted q-value < 0.05 were considered statistically significant (Huggett et al., 2023). Visualizations for the differentially spliced clusters and genes were created using the leafviz extension in leafCutter. GO enrichment analysis for genes containing differentially alternatively spliced intron clusters was performed using the enrichGO function with GO "BP" ontology terms in clusterProfiler.

3.3.13 Integrative gene locus analysis

Integrated visualization of functional genomics data was performed using SparK v.2.6.2 (Kurtenbach & Harbour, 2019). Summary statistics from a meta-analysis GWAS of cocaine dependence (CD) in an EA population ($N = 6378$) (Cabana-Domínguez et al., 2019) was used to cover SNPs that are associated with CD. The EWAS summary statistics from (Poisel et al., 2023) were used as the DNAm dataset. To prioritize the association results for visualization, SNPs and CpG sites with nominal significant association *p*-value ($p < 0.05$) were filtered from the GWAS and EWAS results. ChIP-seq datasets for different activating and repressing chromatin marks were downloaded from ENCODE ("An integrated encyclopedia of DNA elements in the human genome," 2012) as deposited in the Human Reference Epigenome Matrix for dorsolateral PFC in males: ENCFF241REN (H3K4me1), ENCFF752EVS (H3K4me3), ENCFF866IWY (H3K27ac), ENCFF149DDW (H3K36me3), ENCFF784SSN (H3K9me3), ENCFF167ASN (H3K27me3). BigWig files were converted to BedGraph using the UCSC bigWigToBedGraph tool. Bam-files from the RNA-seq analysis were indexed using samtools v.1.5 (Li et al., 2009) and then converted to BedGraph using the bamCoverage function from deeptools v.3.5.3 (Ramírez et al., 2016). For the genetic dataset, we performed an additional gene-based association analysis using Multi-marker Analysis of GenoMic Annotation (MAGMA) (de Leeuw et al., 2015). Here, we aimed to quantify the combined association of all SNPs annotated to a gene of interest with CUD as the phenotype.

3.3.14 Multi-omics factor analysis

Multi-omics factor analysis (MOFA) (Argelaguet et al., 2018) was used to jointly analyze the DNAm and gene expression datasets in BA9 aiming for the identification of CUD-associated factors. The factor analysis framework enables an improved characterization of gene and pathway alterations across different omics datasets by investigating the contribution of each omics view such as DNAm or gene expression to a learned factor. Downstream analyses such as GSEA enable the analysis of biological functions that are associated with a factor based on the factor loading of features such as genes that contribute to the biological pathway. As the DNAm input dataset for MOFA (R package v.1.3.1), we used methylation M-values from the 20,000 most variant promoter CpG sites (TSS200 and TSS1500 annotations) under the assumption of their prominent role in regulating transcription levels of nearby genes. Methylation data was extracted for the individuals that also had expression data available ($N = 25$). For the expression dataset, we used normalized and variance stabilization transformed counts from the 20,000 most variant genes to obtain an equal number of features in each view. The MOFA model was trained on the matched DNAm and expression data from $N = 25$ individuals using default model options with a total of 10 factors and the training options convergence_mode = "slow", seed = 42, and maxiter = 10 000. Association of factors with phenotypes was evaluated using the correlate_factors_with_covariates function. GSEA was performed on negative and positive weights individually using the run_enrichment function based on the c5.go.bp.v2023.1.Hs.symbols.gmt gene set reference file from MSigDB (Subramanian et al., 2005). Functional characterization of DNAm weights was performed by subsetting the top 2.5% of CpG sites from both sides of the weight distribution on factor 9 resulting in $N = 500$ CpG sites with strongest positive and negative weights on factor 9, respectively. Next, GO enrichment analysis was performed in missMethyl v.1.33.1 (Phipson et al., 2016) using the full set of $N = 20{,}000$ CpG sites as background.

3.3.15 GO enrichment analysis of CUD-associated gene sets

Convergence of CUD association signals at the pathway level was evaluated by pathway enrichment analysis for GO terms using the enrichGO function on the GO "BP" ontology in the compareCluster functionality of clusterProfiler. A total of 10 gene lists were included in the input dataset: 1) CUD-associated CpG sites ($N = 394$, $p > 0.001$) from the EWAS of CUD (Poisel et al., 2023), genes in the CUD-associated WGCNA methylation modules 2) blue ($N = 9201$), 3) steelblue ($N = 390$), 4) brown ($N = 5268$), 5) brown4 ($N = 205$), 6) nominally significant DEGs ($N = 1057$, $p < 0.05$), 7) genes in the CUD-associated WGCNA expression module yellow ($N = 2517$), 8) AS genes ($N = 98$, $q < 0.05$), 9) MOFA methylation weights factor 9 ($N = 983$ genes based on the 2.5 and 97.5 percentiles of the weight

distribution for CpG sites), and 10) MOFA methylation weights factor 9 ($N = 1000$ genes based on the 2.5 and 97.5 percentiles of the weight distribution for genes). Pathways remaining statistically significant after FDR correction ($q < 0.05$) were displayed in an enrichment map with a pie plot visualization scheme for GO terms that were repeatedly identified for the different gene lists.

3.4 Results

3.4.1 Sociodemographic characteristics and cell type composition estimation in the postmortem brain tissue cohort

We first assessed the phenotypic similarities between CUD cases and controls. No significant differences were observed regarding the pH value of the brain, postmortem interval (PMI), RNA integrity number (RIN), and occurrence of comorbid depressive and alcohol use disorders (Table 3.1). A significant difference was detected for suicide that was the most frequent cause of death in CUD individuals ($p = 7.01e\text{-}03$, Table 3.1).

Table 3.1 – Demographic data for the N=25 Brodmann Area 9 postmortem human brain tissue donors with RNA-seq data
SD, standard deviation; PMI, post-mortem interval; RIN, RNA integrity number; MDD, major depressive disorder; NOS, depressive disorder not otherwise specified; p-value, derived from CUD/no CUD comparison using a t-test for continuous and chi-squared test for categorical variables.

Variable	**CUD**	**No CUD**	**p-value**
N	13	12	-
Age (years(SD))	44.9 (11.3)	52.6 (10.7)	0.10
Sex (Male/Female)	13/0	12/0	-
pH (SD)	6.49 (0.27)	6.28 (0.22)	0.05
PMI (hours(SD))	55.63 (23.00)	55.90 (20.91)	0.98
RIN (SD)	7.63 (0.79)	7.36 (1.18)	0.51
Depressive disorder (MDD, NOS), yes (%)	5 (38.5)	2 (16.7)	0.44
Alcohol use disorder, yes (%)	3 (23.1)	2 (16.7)	1.00
Cause of death			
Accidental	2 (15.4)	4 (33.3)	0.56
Natural	2 (15.4)	7 (58.3)	0.07
Suicide	9 (69.2)	1 (8.33)	**7.01e-03**
Toxicology at death			
Cocaine or metabolites at death, yes (%)	10 (76.9)	0 (0)	**4.41e-04**
Alcohol, yes (%)	4 (30.7)	1 (8.33)	0.37
Opioids, yes(%)	1 (7.69)	0 (0)	1.00
Methamphetamine and metabolites, yes(%)	2 (15.4)	0 (0)	0.50
Cannabinoids, yes(%)	1 (7.69)	1 (8.33)	1.00
Benzodiazepines/GHB, yes(%)	1 (7.69)	0 (0)	1.00
Antidepressants, yes(%)	2 (15.4)	4 (33.3)	0.56
Antipsychotics, yes(%)	1 (7.69)	1 (8.33)	1.00

We further investigated the variance partition of potential covariates in the RNA-seq dataset (Supplementary Fig. S3.1a) and found age, PMI, brain pH, and RIN to

be associated with gene expression levels, and hence included them as covariates in further analyses. To explore whether major cell type composition could affect analysis results, we performed a cell type deconvolution analysis using CIBERSORT based on human PFC major cell type marker gene signatures (Supplementary Fig. S3.1b, Supplementary Table S3.2a). No significant differences in the distribution of major cell types such as astrocytes, oligodentrocytes, microglia, neurons and others were detected between samples from individuals with and without CUD as all 95% high-density intervals from the Bayesian estimation contained 0 (Supplementary Table S3.2b).

3.4.2 Transcriptome-wide differential gene expression patterns in CUD are related to synaptic signaling, ion transport, and inflammatory processes

The transcriptome-wide analysis of differential expression in BA9 revealed a total of $N = 1057$ DEGs associated with CUD ($p < 0.05$). Of these, $N = 378$ were upregulated and $N = 679$ were downregulated (Fig. 3.1a, Supplementary Table S3.3a, b). After adjustment for multiple testing, *ZFAND2A* (*Zinc Finger AN1-Type Containing 2A*, log2FC = 0.43, $p = 1.98e\text{-}06$, $q = 0.04$), remained significantly upregulated in individuals with CUD (5% FDR). Results were stable in a sensitivity analysis when AUD and MDD status were included as additional covariates in differential expression testing (Supplementary Fig. S3.1c). To evaluate whether BA9 DEGs are significantly enriched within cell type-specific genes of the human PFC, we performed an overlap analysis, using the same set of major brain cell type marker genes as in the cell type composition analysis. Upregulated DEGs were significantly enriched for neuron marker genes exclusively, whereas downregulated DEGs were significantly enriched in markers of non-neuronal cell types such as astrocytes, endothelial cells, and oligodendrocytes (Fig. 3.1b).

Next, we were interested in the biological functions related to the identified CUD-associated DEGs. After adjusting for multiple testing, we detected $N = 276$ statistically significant GO terms for positive GSEA normalized enrichment scores (NES, Supplementary Table S3.4) and $N = 782$ significant GO terms for negative NES (Supplementary Table S3.5). Among significantly enriched pathways, the largest positive NES was detected for "vesicle-mediated transport in synapse" (NES = 2.63, $q = 5.31e\text{-}15$), whereas "superoxide metabolic process" (NES = −2.43, $q = 2.08e\text{-}06$) was the top finding with negative NES. To identify functional modules of pathways consisting of multiple GO terms related to similar biological functions, we created an enrichment map (emap) visualization based on the significant findings from GSEA. For GO terms with negative NES, we detected one large cluster related to inflammatory and immune signaling and several smaller clusters consisting of pathways involved in angiogenesis, extracellular matrix (ECM) organization, and gliogenesis (Fig. 3.1c). Two major clusters

emerged for pathways with positive NES. The first was related to neurotransmission and synaptic signaling whereas the second cluster consisted of GO terms involved in transmembrane transporter activity (Fig. 3.1d).

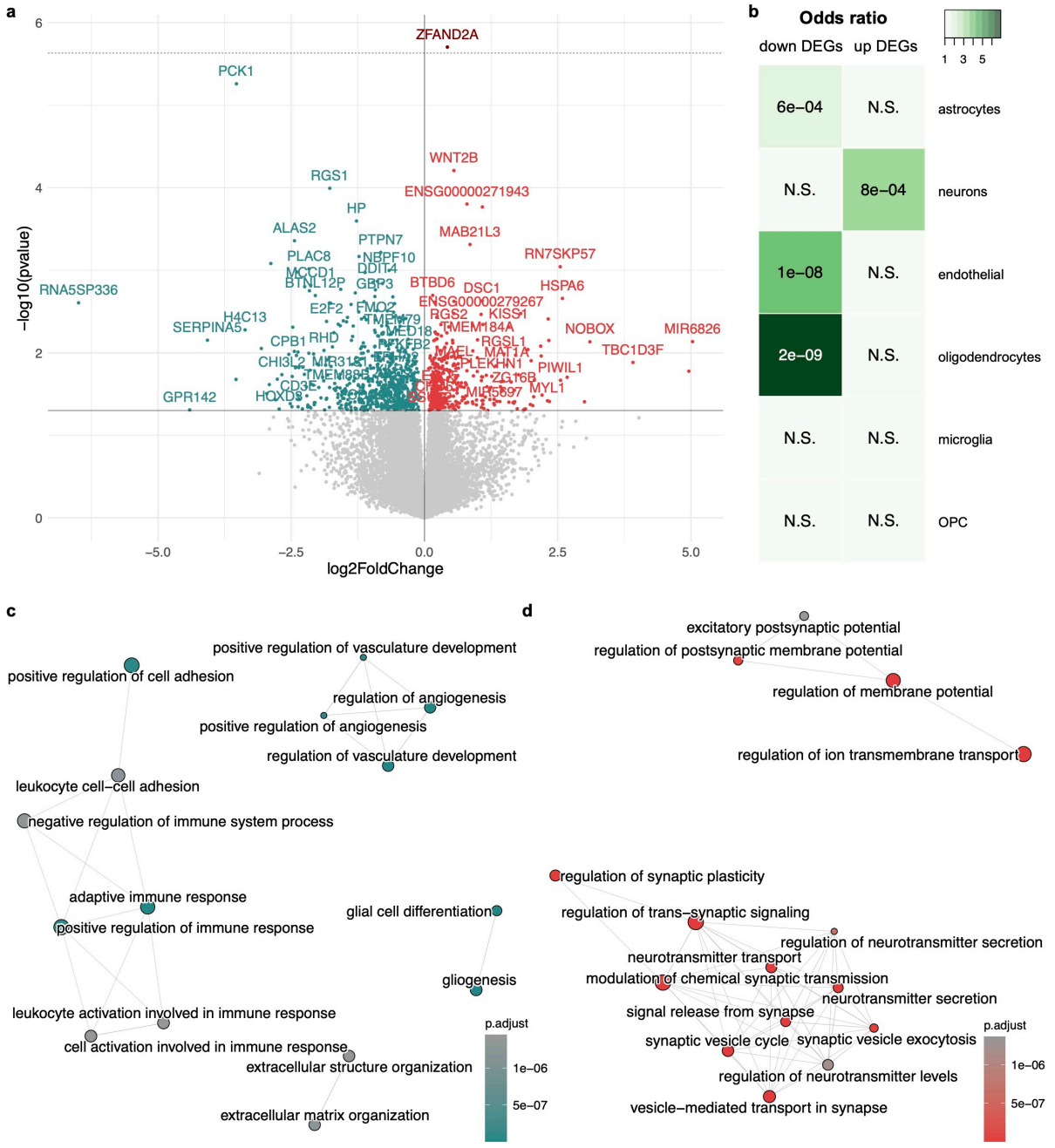


Figure 3.1 - Differential expression analysis in CUD suggests synaptic signaling and immunological alterations in Brodmann Area 9.
a Volcano plot of the differential expression (DE) analysis showing the N = 378 upregulated (red) and N = 679 downregulated genes (blue) at nominal significance ($p < 0.05$). Solid black line indicates nominal significance ($p < 0.05$), dashed gray line indicates transcriptome-wide significance (FDR $q < 0.05$). **b** Results of the overlap analysis for upregulated (up) and downregulated (down) DEGs among cell type-specific marker genes. Green color depicts the odds ratio (OR) of overlap, p-values inside the panels indicates significance of overlap based on Fisher-Test. Gene-set enrichment analysis (GSEA) was performed for the DEGs in BA9 ranked by the Wald test statistic from DESeq2. Statistically significant results ($q<0.05$) from GSEA separated by **c** negative and **d**

positive normalized enrichment scores (NES) are shown in an enrichment map visualization. N.S. = not significant, OPC = oligodendrocyte progenitor cell.

3.4.3 Network analysis highlights fatty acid metabolism and morphogenesis processes in CUD

We next performed WGCNA to investigate gene co-expression patterns in CUD and detected a total of $N = 27$ co-expression modules (Supplementary Fig. S3.2a, 3.2b). Co-expression module yellow was significantly correlated with CUD ($r = -0.47$, $p = 0.02$) while no significant association with other known covariates was observed (Supplementary Fig. S3.2a). Module yellow consisted of $N = 2517$ genes and module membership was highly correlated with gene significance for CUD ($r = 0.61$, $p <$ 1e-200, Supplementary Fig. S3.2c, Table S3.6). GO enrichment analysis for module yellow genes revealed $N = 519$ statistically significant GO terms after multiple testing correction (Supplementary Table S3.7). Strongest associations were detected for "small molecule catabolic process" ($q =$ 2.22e-11), and more specifically, "carboxylic acid catabolic process" ($q =$ 1.59e-09, Supplementary Fig. S3.2d). After clustering the significant terms, a prominent GO term cluster related to fatty acid metabolism was detected, while another cluster was related to organ developmental and morphogenesis processes. To further characterize WGCNA expression module yellow, we generated a protein-protein interaction (PPI) network based on module hub genes and identified APOE ($N = 9$ edges), ERBB2 ($N = 8$ edges), ALDH7A1 ($N = 7$ edges), PPARA ($N = 7$ edges), and TLR4 ($N = 7$ edges) as the most strongly connected nodes in the PPI network (Supplementary Fig. S3.3, Supplementary Table S3.6).

3.4.4 Genes with alternative splicing events in CUD are involved in cell junction formation and the morphogenesis of axons and dendrites

To investigate alternative splicing in CUD and its potential relevance for contributing to altered neurobiological functions in the brain, we performed a differential alternative splicing analysis using LeafCutter (Li et al., 2018). After multiple testing correction, we identified a total of $N = 108$ differentially spliced intron clusters in BA9 (FDR < 0.05, Fig. 3.2a, Supplementary Tables S3.8–3.10). These clusters were distributed among $N = 98$ genes that we further denote as alternatively spliced (AS) genes. One of the top findings in our AS analysis of CUD was *BIN1* (*Bridging Integrator 1*, $q =$ 7.8e-04, Fig. 3.2b) previously identified as a conserved AS genes in other substance use disorders. We next investigated the biological pathways enriched for alternative splicing events based on our list of AS genes. Statistically significant enrichment after multiple testing correction was detected for $N = 15$ GO terms (Supplementary Table S3.10). Strongest enrichment was found for GO terms "cell junction assembly" and "neuron projection extension" (both $q =$ 3.62e-03). In the emap visualization of enriched GO terms with a more lenient threshold of 25% FDR (q < 0.25), we detected a well-connected cluster

containing GO terms related to cellular growth and cell-cell junction development, while also more brain-specific processes such as myelination and the extension of axons and dendrites were found (Fig. 3.2c). While differential alternative splicing itself contributes to altered biological functions by inducing different abundances of transcript isoforms, this effect might be potentiated by differential gene expression. We thus investigated the overlap of AS and DEGs in CUD and identified 8 genes that were differentially spliced and differentially expressed in BA9: *ITPKB, CPLX1, HLA-F, INPP5E, GALNT8, IGFBP6, ZBTB4*, and *BCAT2* (Fig. 3.2d).

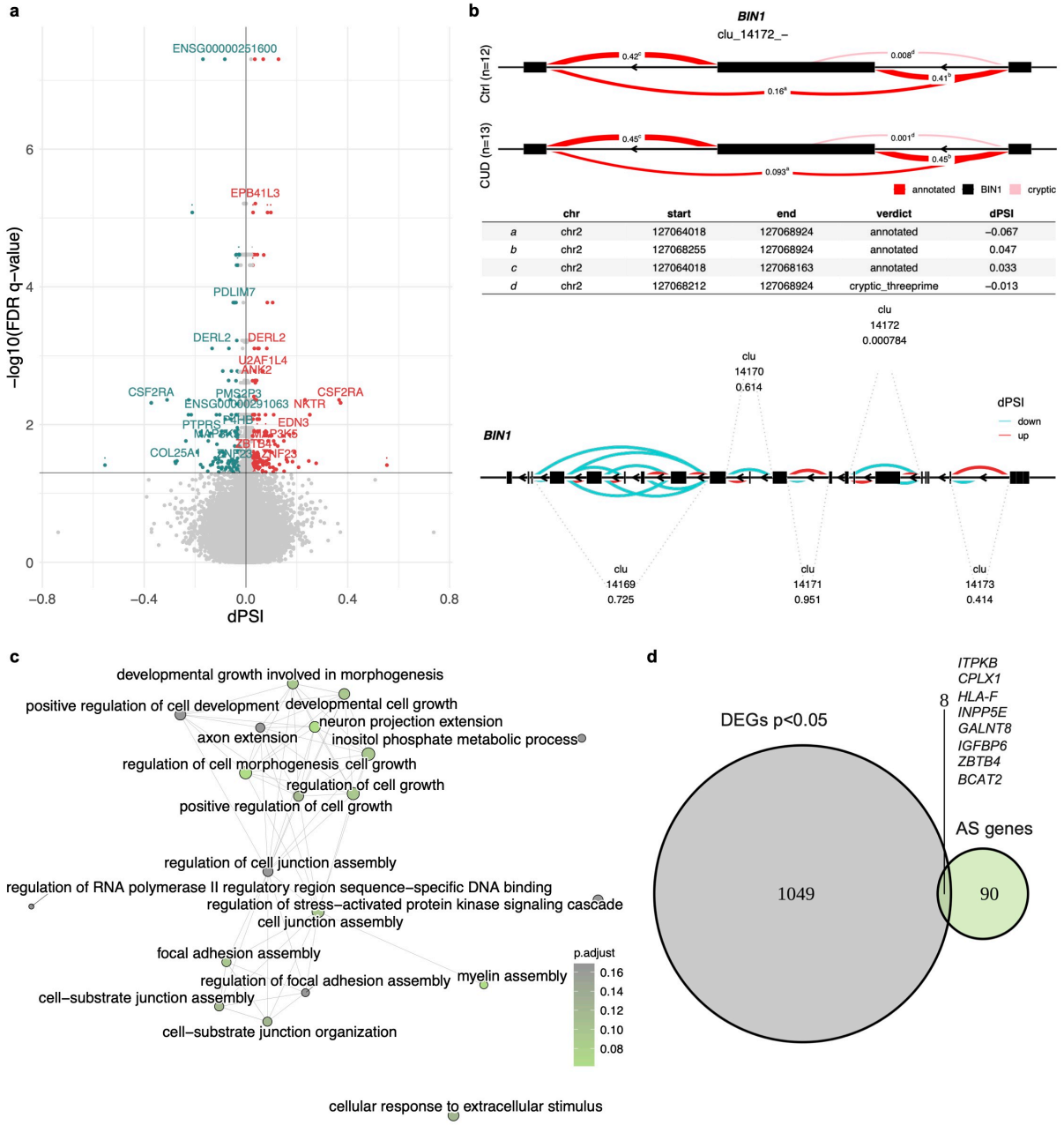


Figure 3.2 – Differential alternatively spliced genes in CUD are related to neuron morphogenetic processes.
a Volcano plot of the differential alternative splicing (AS) results in Brodmann Area 9 (BA9). Statistically significant intron clusters (N = 108) identified by LeafCutter ($q < 0.05$) were annotated by gene name while dots represent individual introns of an intron cluster. Introns highlighted in

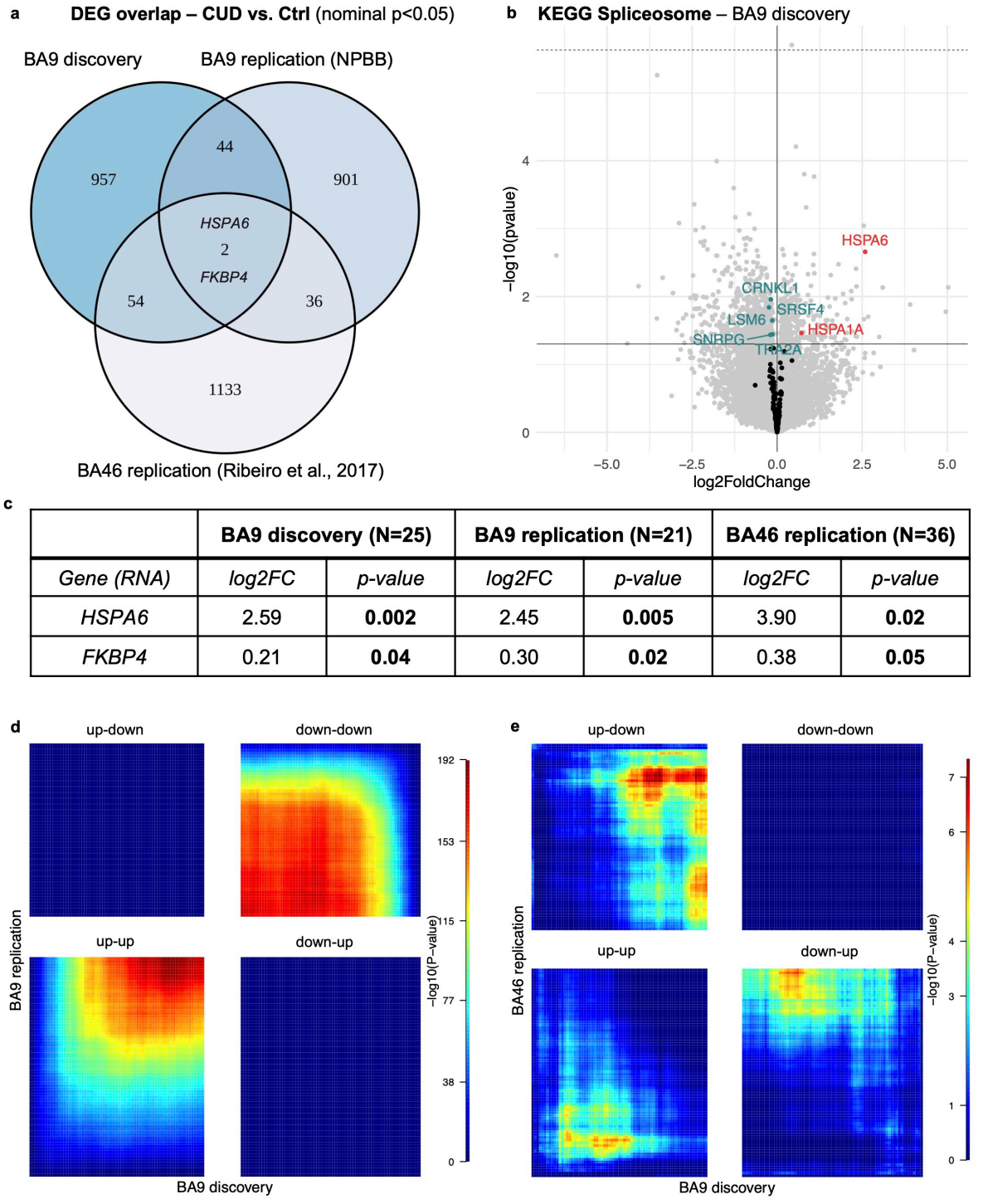
red (dPSI>0) are more abundant in CUD while introns highlighted in blue (dPSI<0) are less abundant in CUD. **b** Results of the differential AS analysis at the cluster and gene level for one of the top findings, an intron cluster (clu_14172_-) in the Bridging Integrator 1 (BIN1) gene. Upper panel: visualization of BIN1 exons and introns with percent spliced in (PSI) measures related to the significant cluster clu_14172_-. The table indicates delta percent spliced in (dPSI) values from the CUD vs. Ctrl comparison. Lower panel: gene-level summary of all intron clusters detected in BIN1. FDR q-values are shown below cluster names. **c** GO enrichment analysis for the N = 98 AS genes harboring differentially statistically significant (q < 0.05) intron clusters in CUD. **d** Overlap of findings from differential expression (DE) analysis (N = 1057 DEGs at p < 0.05) and differential AS analysis.

3.4.5 Replication analysis in independent cohorts reveals *FKBP4* and *HSPA6* as conserved DEGs in CUD

To evaluate the potential replication of CUD-associated DEGs in other RNA-seq datasets of human PFC, we performed an overlap analysis of nominally significant DEGs ($p < 0.05$) across studies. CUD-associated differential expression testing was performed in two independent replication datasets, the first originating from BA9 (BA9 replication, bulk RNA-seq) and the second from BA46 (BA46 replication, neuron-specific RNA-seq, Supplementary Fig. S3.4). Two genes, *HSPA6* and *FKBP4*, were shared upregulated DEGs at nominal significance and showed comparable effect sizes (log2FC) in CUD across all three PFC datasets (Fig. 3.3a, c). As *HSPA6* is a spliceosome-associated gene with conserved differential expression across datasets, we performed a look-up of genes related to the KEGG Spliceosome pathway (hsa03040) in DE results from our discovery cohort (Fig. 3.3b). Here, we aimed to address the hypothesis of spliceosomal differential gene expression as a potential mechanism for splicing alterations in CUD (Huggett & Stallings, 2020). *HSPA6* was the spliceosome-associated gene showing strongest CUD-associated expression changes in BA9 (log2FC = 2.59, $p = 0.002$). We detected six additional spliceosome-associated genes that were among nominally significant DEGs: *HSPA1A* (log2FC = 0.71, $p = 0.034$), *CRNKL1* (log2FC = −0.19, $p = 0.011$), *LSM6* (log2FC = −0.24, $p = 0.014$), *SRSF4*(log2FC = −0.14, $p = 0.022$), *SNRPG* (log 2FC = −0.20, $p = 0.037$), and *TRA2A* (log2FC = −0.14, $p = 0.036$). Using rank-rank hypergeometric overlap (RRHO) visualization for a more unbiased evaluation of convergent and divergent gene expression patterns across studies, we found strong convergent overlap between the BA9 discovery and BA9 replication datasets indicating similar patterns of CUD-associated expression deregulation (Fig. 3.3d). In the comparison with neuron-specific expression data from BA46, we found prominent divergent gene expression patterns between datasets, while convergent expression patterns were detected in the shared upregulated genes across studies (Fig. 3.3e).

3.4.6 Drug repositioning analysis highlights glucocorticoid receptor targeting drugs to reverse the CUD gene expression profile

To evaluate the potential use of small molecule drugs to revert the gene expression pattern of CUD, we performed drug repositioning analysis based on the L1000 assay as implemented in CMap (Supplementary Fig. S3.5a), using the top 150 up- and downregulated genes as input (Supplementary Table S3.11). Among the results with negative normalized connectivity score (NCS), i.e. perturbagens that revert the DE profile in CUD, the most significant finding for small molecule drugs after multiple testing correction was the glucocorticoid receptor agonist medrysone (NCS = −1.78, q = 2.2e-16, Supplementary Fig. S3.5b). Glucocorticoid receptor agonists were the only FDR-significant perturbagen class overrepresented among CMap GSEA results (Supplementary Fig. S3.5c). When we further investigated connectivity scores for all glucocorticoid receptor targeting drugs including agonists and antagonists in CMap, we found exclusively significant negative connectivity scores (Supplementary Fig. S3.5d) suggesting glucocorticoid receptor targeting molecules as potential pharmacological drugs to revert the CUD expression changes in BA9. In line with this finding, the biological pathway "response to glucocorticoid" (NES = −1.54, q = 0.019, Supplementary Table S3.5) was among the FDR-significant pathways with negative NES in the GSEA analysis of DEGs from BA9.



	BA9 discovery (N=25)		BA9 replication (N=21)		BA46 replication (N=36)	
Gene (RNA)	*log2FC*	*p-value*	*log2FC*	*p-value*	*log2FC*	*p-value*
HSPA6	2.59	**0.002**	2.45	**0.005**	3.90	**0.02**
FKBP4	0.21	**0.04**	0.30	**0.02**	0.38	**0.05**

Figure 3.3 - Replication analysis of CUD associated transcriptomic alterations in independent datasets. **a** Overlap of nominally significant ($p < 0.05$) differentially expressed genes across datasets reveals two shared DEGs, HSPA6 and FKBP4. The replication datasets were based on N = 21 BA9 samples from the National PTSD Brain Bank (NPBB) and neuronal-specific transcriptomic data of N = 36 BA46 samples available under GEO accession number GSE99349 (BA46 replication). **b** HSPA6 is the strongest spliceosome-associated DEG in BA9. **c** Results of the look- up analysis for shared DEGs HSPA6 and FKBP4 - log2-fold change and p-value: association p-value from the DESeq2-based differential expression results. Significant associations are highlighted in bold. Rank-rank hypergeometric overlap (RRHO) visualization for **d** the BA9 replication dataset and **e** the neuronal-specific BA46 dataset indicating convergent and divergent expression patterns across studies using full differential expression statistics as the input datasets. Color scale represents -

log10(p) of the hypergeometric testing procedure in RRHO. Convergent expression across datasets: up-up and down-down, divergent: up-down and down-up.

3.4.7 Findings of the integrated analysis of DNA methylation and gene expression data converge at the gene and pathway levels

As DNA methylation data was available and previously analyzed for the same cohort in BA9, we next aimed to integrate findings from the epigenome-wide and transcriptome-wide studies on the gene-level, applied multi-omics factor analysis, and performed an integrative functional GO-term enrichment analysis across all –omics layers.

3.4.7.1 Gene-level integration of epigenomic, transcriptomic and splicing alterations highlights ZBTB4 and INPP5E in CUD

Of the overlapping genes between the differential methylation, expression, and alternative splicing analyses, two genes were consistently altered across all the investigated molecular views in BA9: *ZBTB4 (Zinc Finger And BTB Domain Containing 4)* and *INPP5E (Inositol polyphosphate-5-phosphatase E)* (Fig. 3.4a, Supplementary Table S3.12). Both genes were characterized by a hypomethylated CpG site and increased transcript levels in CUD (Fig. 3.4b). For *ZBTB4*, the strongest association for a CpG site was found for cg03443505 (chr17:7387573, $\beta = -0.84$, $p = 1.01e\text{-}05$). *ZBTB4* was upregulated with a log2FC of 0.08 ($p = 0.015$) and it contained the differentially spliced intron cluster *chr17:clu_10246_-* ($q = 0.028$). The strongest association for CpG differential methylation in the *INPP5E* gene was found for cg18558462 (chr9:139334381, $\beta = -0.93$, $p = 8.55e\text{-}03$). It was differentially expressed with log2FC of 0.17 ($p = 0.025$) and intron cluster *chr9:clu_25078_-* was differentially alternatively spliced ($q = 0.015$). In the replication datasets, we detected conserved transcript upregulation of *ZBTB4* (BA9 replication, log2FC = 0.12, pval=0.21; BA46, log2FC = 0.12, pval=0.09) and *INPP5E* (BA9 replication, log2FC = 0.11, pval=0.47; BA46, log2FC = 0.08, pval=0.49), however not statistically significant.

To deeper characterize the *ZBTB4* and *INPP5E* gene loci in BA9 and specifically in the context of CUD, we performed an integrative gene locus visualization approach by combining GWAS, EWAS, alternative splicing, and RNA-seq results for CUD with ENCODE ChIP-seq reference data from human dorsolateral prefrontal cortex. ChIP-seq data confirmed the presence of activating chromatin marks at promoter (H3K4me4, H3K27ac) and gene body regions (H3K36me3) at *ZBTB4* and *INPP5E* gene loci in the human dlPFC. In addition, for *ZBTB4*, multiple nominally significant associations for SNPs and CpG sites were detected that were most prominent within intronic and intergenic regions, while no SNP but CUD-associated CpG sites were identified in the *INPP5E* gene locus (Fig. 3.4c). In line with this, when we quantified the association of genetic variants

with CUD at the gene level using a gene-based association analysis in MAGMA, we detected stronger association for *ZBTB4* with CUD ($Z = 1.74$; $p = 0.04$) compared to *INPP5E* ($Z = -1.14$; $p = 0.87$).

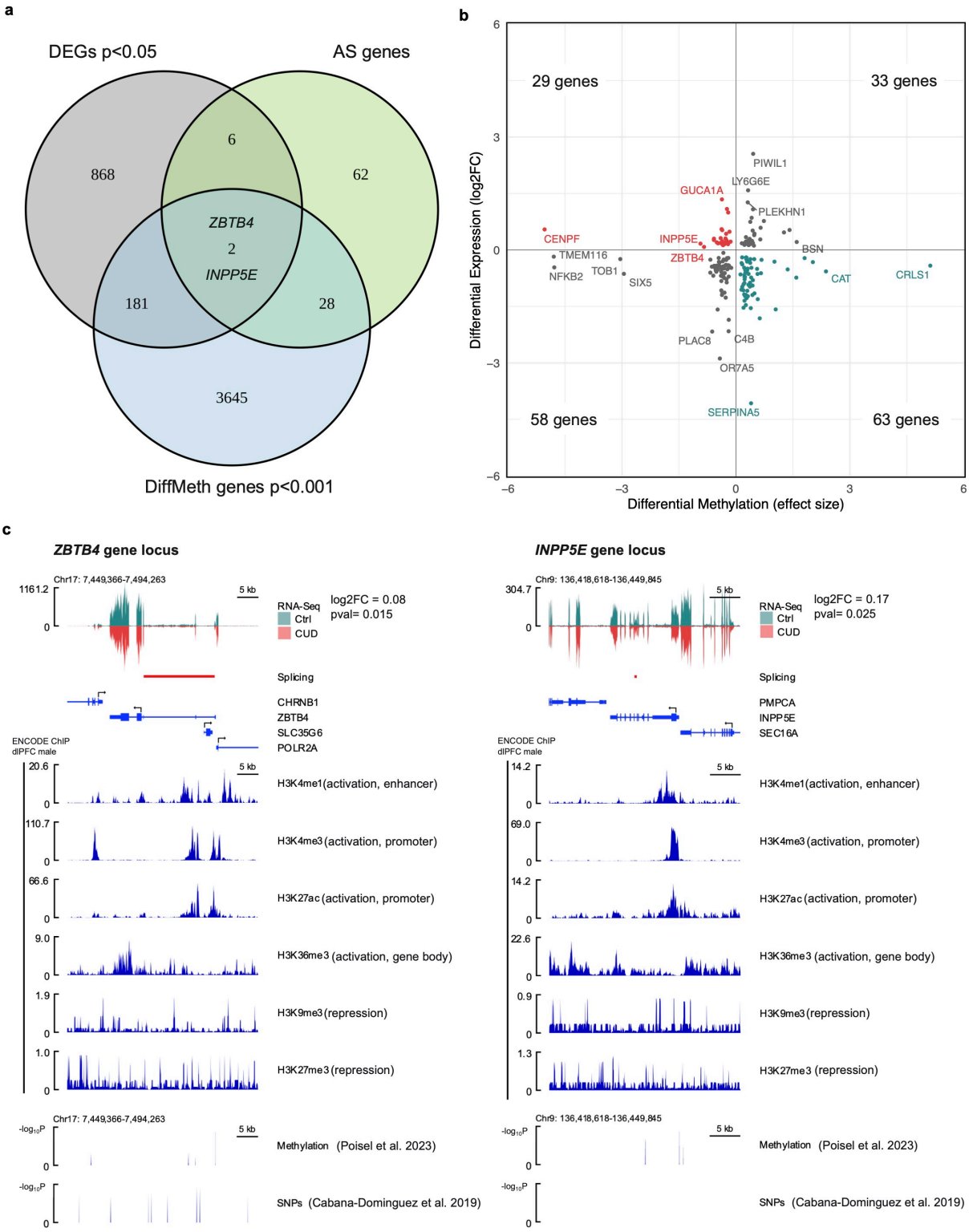


Figure 3.4 - Convergence of DNA methylation, alternative splicing, and gene expression alterations in CUD at the ZBTB4 and INPP5E gene loci. **a** Overlap of differential expression (DE), differential DNA methylation (DiffMeth), and differential alternative splicing analyses suggest two genes, ZBTB4 and INPP5E, where alterations are consistently associated with CUD. **b**

Relationship between DE and DiffMeth genes in Brodmann Area 9 based on log2FC (y-axis) from DE analysis and effect size β from linear regression in the EWAS of CUD (x-axis). For both genes, ZBTB4 and INPP5E (highlighted in red), hypomethylation of the strongest significant CUD-associated CpG site and increased transcript levels are observed. **c** Integrated visualization of functional genomic datasets for ZBTB4 and INPP5E gene loci. CUD-associated genomic variants (SNPs $p < 0.05$ from 849]), CUD-associated CpG sites ($p < 0.05$ from [22]), RNA-seq data and intron clusters ($q < 0.05$) from the present study were visualized together with ENCODE ChIP-seq data for different chromatin marks in human dorsolateral prefrontal cortex.

3.4.7.2 Multi-omics factor analysis confirms cell junction, synaptic signaling, and neurogenesis as important biological processes in CUD

The integration of DNA methylation and gene expression data as described above was based on results of the EWAS, DE, and AS analyses depicting one possible way of integrating multiple omics datasets. In addition, multi-omics analysis tools such as MOFA enable an integrated analysis of omics datasets in a single statistical framework. Using MOFA on our DNA methylation and gene expression data from BA9, we identified one factor representation of the multi-omics dataset (factor 9) that was significantly correlated with CUD ($r = -0.48$, $p = 0.02$) and age ($r = 0.47$, $p = 0.02$, Supplementary Fig. S3.6a, b). Factor 9 displayed significantly smaller factor values in CUD cases compared to individuals without CUD in a Wilcoxon test ($p = 0.02$, Supplementary Fig. S3.6c). When we extracted the CpG sites with the strongest weights on factor 9, cg23859635 annotated to *MTA3* was the CpG site with the strongest positive weight on factor 9 ($w = 0.31$), while cg24621354 in the gene *TES* displayed the strongest negative weight ($w = -0.33$, Supplementary Fig. S3.6d, Supplementary Table S3.13). In the gene expression dataset, the small GTPase *RAB6A* had the strongest positive weight ($w = 0.07$), while *HIVEP2* had the strongest negative weight ($w = -0.05$) on factor 9 (Supplementary Fig. S3.6e, Table S3.13). Results of a GSEA on negative expression weights on factor 9 revealed FDR-significant ($q < 0.05$) enrichment for synaptic signaling, cell junction organization, and neurogenesis pathways, confirming the results from the previous analyses. In contrast, GSEA on positive expression weights revealed enrichment for cellular respiration and small molecule metabolic processes (Supplementary Fig. S3.6f, Supplementary Table S3.14). When we used missMethyl to investigate the biological pathways for DNA methylation features with strong weights on factor 9, we detected enrichment for similar biological pathways as in the analysis of expression features. While none of the enrichment results remained FDR-significant after multiple testing correction, strongest enrichment for CpG sites with negative weights on factor 9 was detected for intracellular calcium concentration regulation and synaptic vesicle related processes (Supplementary Fig. S3.6g, Supplementary Table S3.15). The pathways showing the strongest enrichment for the positive weight CpG sites were related to monocarboxylic acid and specifically, lactate transmembrane transporter activity, and ER stress pathways (Supplementary Fig. S3.6h, Supplementary Table S3.15).

3.4.7.3 Integrative functional analysis reveals pathway modules related to neurotransmission, cell differentiation, cell junction organization, and fatty acid metabolism

In an integrative functional analysis approach, we used all available information from our study on DNA methylation, gene expression, and alternative splicing alterations in CUD to identify potential convergence of association results at the pathway level in BA9. We thus performed a GO enrichment analysis based on 10 curated lists with CUD-associated genes derived from EWAS, DE analysis, alternative splicing analysis, WGCNA modules based on DNA methylation and expression data, and MOFA (Supplementary Table S3.16). In the enrichment map for GO terms, we identified several functional modules where the same biological pathway was detected for multiple gene lists at FDR-adjusted statistical significance ($q < 0.05$) indicating convergence of the results from different analysis approaches (Supplementary Fig. S3.7). The two largest functional modules (FM) contained pathways involved in neurotransmission and synaptic signaling (FM1), while FM2 was enriched for neuron and glial cell differentiation, growth, and morphogenesis processes. Two further prominent modules were related to synapse and cell junction organization (FM3) and fatty acid metabolism (FM4).

3.5 Discussion

By applying a multi-omics data integration approach on DNA methylation and gene expression data from postmortem human brain tissue we aimed for a deeper understanding of the neurobiology of CUD in the human prefrontal cortex. At the gene level, our differential expression analysis suggests two candidates, *FKBP4* and *HSPA6*, which were replicated as nominally significant findings in two independent cohorts. In addition, our multi-omics analyses highlight *ZBTB4* and *INPP5E*, that were consistently altered across omics analyses in BA9 and displayed consistent upregulation patterns in independent replication datasets. At the pathway level, we found converging evidence for CUD-associated DNAm and transcriptional alterations that were related to neurotransmission, fatty acid metabolism, and changes in neuronal morphology.

Analysis of the transcriptome in BA9 revealed *ZFAND2A* as the DEG showing the strongest association with CUD. *ZFAND2A* is a canonical heat shock gene in humans encoding a zinc-finger containing protein that is involved in the regulation of proteasomal protein degradation (Lee et al., 2018; Rossi et al., 2014). It was further identified as a DEG in a study on transcriptomic signatures of Alzheimer's disease (AD) (Shippy & Ulland, 2022). Another AD-related finding emerged in co-expression network analysis. APOE showed the strongest connectivity in the PPI network for module yellow genes and has been intensively characterized due to its association with age of onset in AD (Yamazaki et al., 2019). While SUDs and

neurodegenerative disorders such as AD depict different neuropsychiatric disorders based on the current understanding of disease mechanisms, CUD and AD share brain atrophic changes as a clinical symptom (Ersche, Jones, et al., 2013) and our results suggest that there might be shared molecular mechanisms involved.

Previous studies have identified differential alternative splicing in AUD (Huggett et al., 2023) and OUD (Huggett et al., 2022) in the human brain, however, RNA splicing alterations have not been characterized in human CUD so far. In the differential alternative splicing analysis, we found $N = 98$ statistically significant genes containing AS intron clusters. Interestingly, among our top findings, we found *Bridging Integrator 1* (*BIN1*) for which differential alternative splicing in the brain has been described in OUD. *BIN1* was the only differential AS gene in OUD that was conserved across all investigated brain regions; dlPFC, NAc, and midbrain (Huggett et al., 2022). Further, AS events in *Bin1* were identified in the mouse brain in a study on splicing alterations associated with cocaine self-administration (Xu et al., 2021). As dendrite and axon morphogenesis processes were among the enrichment results for AS genes in BA9, we hypothesize that AS is directly related to neuroplastic changes in the CUD brain. Mechanistically, AS processes change the abundance of transcript isoforms with different biological functions that might contribute to the neuroadaptations in CUD. We explored the mechanism of spliceosomal gene expression alterations as a potential contributor to differential AS events in CUD. Exposure to cocaine was previously hypothesized to alter spliceosomal gene expression (Huggett & Stallings, 2020) and our results suggest spliceosomal genes such as *HSPA6* and *HSP1A1* as DEGs in BA9. As spliceosomal gene alterations were also detected in the replication analysis with *HSPA6* as a shared upregulated DE gene across studies, AS events might be an important mechanism in CUD contributing to neurobiological changes in the PFC.

In the last step of the RNA-sequencing analysis, we aimed to address the urgent need for novel pharmacotherapeutic approaches for the treatment of CUD by performing a drug repositioning analysis. We detected glucocorticoid receptor-targeting drugs having consistently negative connections with the CUD expression profile in BA9. In addition, *FKBP4*, an important regulator of glucocorticoid receptor signaling was identified as a conserved upregulated DEG in CUD based on three independent dlPFC datasets. FKBP4 has a key role in the nuclear translocation of the glucocorticoid receptor, as it replaces FKBP5 upon cortisol binding to the receptor complex leading to its nuclear translocation (Davies et al., 2002). Pharmacological targeting of glucocorticoid receptor signaling was tested in rodent models of cocaine addiction (Deroche-Gamonet et al., 2003; Evan et al., 2013; John et al., 1998). Reduced behavioral response to cocaine was observed when glucocorticoid receptor antagonists such as mifepristone were applied

(Deroche-Gamonet et al., 2003). In contrast, corticosterone was shown to promote cocaine intake in rats (Evan et al., 2013; John et al., 1998). In the drug repositioning analysis, results for glucocorticoid receptor agonists were more prominent compared to antagonists which appears to be in conflict with previous literature. However, glucocorticoid receptor antagonists such as mifepristone also displayed significant negative connectivity scores with the BA9 expression signature supporting previous findings. Further, synthetic glucocorticoid receptor agonists such as dexamethasone were shown to impair cocaine self-administration in rats (John et al., 1998) indicating a more complex relationship between the endogenous glucocorticoid system and exogenously applied glucocorticoid receptor targeting drugs. We thus suggest that glucocorticoid receptor targeting drugs should be further investigated for their potential use as a pharmacotherapy in CUD.

Using multi-omics data integration, we identified two genes, *ZBTB4* and *INPP5E*, for which CUD-associated alterations were consistently detected across DNAm, gene expression, and alternative splicing analyses. Both genes contained a hypomethylated CpG site, stronger transcript expression was found in individuals with CUD, and significant differentially spliced intron clusters were identified. Despite being strongly expressed in the brain and most prominently in neurons [67], the role of *ZBTB4* in neuropsychiatric disorders remains poorly understood. However, due to the DNA binding capacity and its role as a transcriptional repressor, *ZBTB4* deregulation in CUD could lead to downstream expression changes of its target genes. Further, protein-protein interaction data suggests interaction of ZBTB4 with the transcription factor PRDM5 as well as with the AP2M1 and AP2A1 subunits of the adapter protein 2 (AP-2) complex that is involved in endocytosis of neurotransmitter receptors in neurons (Guardia et al., 2018; Szklarczyk et al., 2023). The second finding at the gene level, *INPP5E*, encodes a phosphatidylinositol-phosphatase specific to cilia and *INPP5E* mutations were found in Joubert syndrome which is characterized by cerebellar and cerebral malformation (Hardee et al., 2017). A possible link to CUD provide neuronal primary cilia, known as key signaling hubs on somata enriched for G-protein-coupled receptors (GPCRs) (DeMars et al., 2023). As INPP5E is required for proper trafficking of GPCRs along ciliary microtubules (Garcia-Gonzalo et al., 2015), deregulation of *INPP5E* might lead to aberrant ciliary signaling that has recently gained attention in the addiction field: cell type-specific ablation of neuronal primary cilia in mice was shown to affect body weight as well as locomotor response to psychostimulants such as cocaine (Everett et al., 2024) and amphetamine (Ramos et al., 2021). In humans, further studies on INPP5E are required to evaluate its role in SUDs.

Evaluating the convergence of results at the pathway level revealed widespread molecular alterations in synaptic signaling represented by functional module FM1

in the GO enrichment analysis. Our findings are well in line with previous literature that reported on cocaine-associated DNAm and expression changes in genes involved in neurotransmission (Campbell et al., 2021; Gao et al., 2017; Li et al., 2017; Savell et al., 2020; Vaillancourt, Chen, et al., 2021). The observed overrepresentation of neuronal marker genes in the upregulated DEGs together with the non-neuronal marker gene enrichment in the downregulated DEGs further suggests a particular importance of CUD-associated expression changes in altering neurotransmission. In a study on CUD-associated gene expression changes in neuronal nuclei of the human dlPFC (Ribeiro et al., 2017), the authors found a WGCNA co-expression module that was significantly associated with CUD and was enriched for GTPase signaling and neurotransmitter transport that well matches our results in BA9. Neuronal function thus appears to be strongly influenced by altered epigenetic and transcriptional programs in the CUD brain.

Functional modules FM2 and FM3 were related to pathways involved in neuron, synapse, and axon morphogenetic processes. This is supported by literature from animal models of CUD, where alterations in dendritic branches and spine density were observed in the PFC of cocaine self-administering rats (Robinson et al., 2001). Even a single cocaine exposure was sufficient to reduce dendritic spine density in neurons (Caffino et al., 2018). In summary, brain morphological changes depict an interesting link between molecular and behavioral aspects of addiction as neuroplastic changes are the basis of neurocircuit alterations in the SUD brain that are related to compulsive drug-seeking and relapse (Koob & Volkow, 2010).

Another converging finding were metabolic changes related to fatty acid metabolism (FM4). This finding was especially prominent in the CUD-associated WGCNA module yellow where we found a functional module of pathways related to fatty acid metabolism. Further, results from MOFA suggested gene expression changes related to the electron transport chain as another key metabolic pathway alteration. This is supported by findings from animal models of cocaine addiction where a downregulation of glycolysis and oxidative phosphorylation were observed in the brain (Zhou et al., 2011) while fatty acid metabolism genes were upregulated (Pati et al., 2019). It has to be noted that metabolic changes in CUD are most likely not brain-specific but also appear on a systemic level as individuals with CUD were found to have reduced body fat in comparison to a healthy control group (Ersche, Stochl, et al., 2013). To follow up on this finding, future studies should evaluate if interfering with fatty acid metabolism could depict a therapeutic strategy in CUD as a ketogenic diet has been shown to alter the behavioral response to cocaine in rats (Martinez et al., 2019).

There are some limitations that apply to our multi-omics study of CUD. First, depicting an inherent limitation of analyses in human postmortem brain tissue,

our cross-sectional analysis design can only reflect the endpoint of CUD, limiting the identification of dynamic changes in DNAm and gene expression during the disease course. Second, considering the sample size and the few DEGs at transcriptome-wide significance, it remains unclear whether the findings are generalizable to the general population. While we were able to replicate findings at nominal significance, our studies lack statistical power in the main and replication analyses. This highlights the need for collaborative efforts to perform studies in larger and more diverse cohorts, including meta-analyses. Third, while the homogeneity of our sample consisting of only males from EA ancestry is a strength in statistical analysis, sex-specific and ancestry-related molecular signatures of CUD remain an open question. At least in the analysis of the more diverse replication cohorts we were able to show comparable CUD-associated gene expression patterns when compared to our discovery cohort. At the same time, we were not able to include potentially important covariates, such as cocaine at death and cause of death, in our analyses, because of statistical power. Future efforts with larger sample sizes are needed to investigate how these factors influence CUD-related gene expression. Lastly, it cannot be assumed that differential mRNA expression necessarily leads to changes in protein expression. Therefore, it will be crucial for future studies to combine transcriptomics with proteomics.

In summary, our study identifies novel associations with CUD at the gene level, confirms these on the multi-omics level, and suggests differential alternative splicing as an important molecular hallmark of CUD in the human prefrontal cortex. At the same time, our study supports previous findings of synaptic signaling alterations that have been robustly detected when investigating the neurobiological effects of cocaine. We highlight drugs targeting glucocorticoid receptor signaling to be further tested as a treatment for CUD.

3.6 Data availability

Raw methylation data is deposited in the European Genome Phenome Archive (EGA) under accession number EGAS00001006828 (https://ega-archive.org/studies/EGAS00001006826). RNA-sequencing data has been uploaded to the EGA and is available under accession number EGAS50000000150 (https://ega-archive.org/studies/EGAS50000000150).

3.7 Code availability

All original code used for data analysis and figure preparation is available in a GitHub repository (https://github.com/lzillich/BA9_multi_omics_cocaine).

3.8 Acknowledgements

We thank Elisabeth Röbel and Claudia Schäfer-Arnold for their technical assistance.

3.9 Funding

Funding supporting this study was provided by the German Federal Ministry of Education and Research (BMBF) within the e:Med research program SysMedSUDs: "A systems-medicine approach toward distinct and shared resilience and pathological mechanisms of substance use disorders" (01ZX01909 to RS, MR, ACH, and SHW). In addition, by the Deutsche Forschungsgemeinschaft (DFG) through the collaborative research centre TRR265: "Losing and Regaining Control over Drug Intake" (Heinz et al., 2020) (Project ID 402170461 to SHW, RS, ACH and MR), the Hetzler Foundation for Addiction Research (to ACH), the ERA-NET program: Psi-Alc (01EZ1908), Spanish 'Ministerio de Ciencia, Innovación y Universidades' (PID2021-1277760B-I100, to BC), 'Generalitat de Catalunya/AGAUR' (2021-SGR-01093, to BC), ICREA Academia 2021, 'Fundació La Marató de TV3′ (202218-31, to BC) and from 'Ministerio de Sanidad, Servicios Sociales e Igualdad/Plan Nacional Sobre Drogas' (PNSD-2020I042, to NF-C). The project has been carried out using the Mannheim (CIMH) infrastructure of the German Center for Mental Health (DZPG). Open Access funding enabled and organized by Projekt DEAL.

3.10 Author information

These authors jointly supervised this work: Stephanie H. Witt, Lea Zillich.

Authors and Affiliations

Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Eric Zillich, Hanna Belschner, Diana Avetyan, Diego Andrade-Brito, Josef Frank, Marcella Rietschel, Stephanie H. Witt & Lea Zillich

Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

Diego Andrade-Brito, José Jaime Martínez-Magaña & Janitza L. Montalvo-Ortiz

VA CT Healthcare Center, West Haven, CT, USA

Diego Andrade-Brito, José Jaime Martínez-Magaña & Janitza L. Montalvo-Ortiz

McGill Group for Suicide Studies, Douglas Mental Health University Institute, Montreal, QC, Canada

Naguib Mechawar & Gustavo Turecki

Department of Psychiatry, McGill University, Montreal, QC, Canada
Naguib Mechawar & Gustavo Turecki

Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain
Judit Cabana-Domínguez

Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain
Judit Cabana-Domínguez

Department de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, and Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain
Noèlia Fernàndez-Castillo & Bru Cormand

Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III (ISCIII), Madrid, Spain
Noèlia Fernàndez-Castillo & Bru Cormand

Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Barcelona, Catalonia, Spain
Noèlia Fernàndez-Castillo & Bru Cormand

US Department of Veterans Affairs National Center of Posttraumatic Stress Disorder, Clinical Neurosciences Division, West Haven, CT, USA
Janitza L. Montalvo-Ortiz

Institute of Human Genetics, University of Bonn, School of Medicine and University Hospital Bonn, Bonn, Germany
Markus M. Nöthen

Institute of Psychopharmacology, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
Anita C. Hansson & Rainer Spanagel

German Center for Mental Health (DZPG), partner site Mannheim/Heidelberg/Ulm, Mannheim, Germany
Rainer Spanagel, Stephanie H. Witt & Lea Zillich

Center for Innovative Psychiatric and Psychotherapeutic Research, Biobank, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
Stephanie H. Witt

HITBR Hector Institute for Translational Brain Research gGmbH, Mannheim, Germany
Lea Zillich

Contributions
Conceptualization, EZ, LZ, MR, RS, and SHW; resources, GT, NM, MMN, ACH, JC-D, NF-C, BC, and JMO; data curation, EZ, DA, HB, DA-B, JJM-M, and LZ ; methodology, EZ, HB, JF, and LZ; formal analysis, EZ, HB, DA, DA-B, LZ, and JMM-M, original draft writing, EZ, HB; review & editing, EZ, HB, DA, DA-B, JM-M, JF, NM, GT, JC-D, NF-C, BC, JLM-M, ACH, MR, RS, SHW, and LZ; supervision, JF, MR, and RS, JMO, LZ, and SHW; project administration, RS, MMN, ACH, MR and SHW, funding, RS, ACH, MR, and SHW.

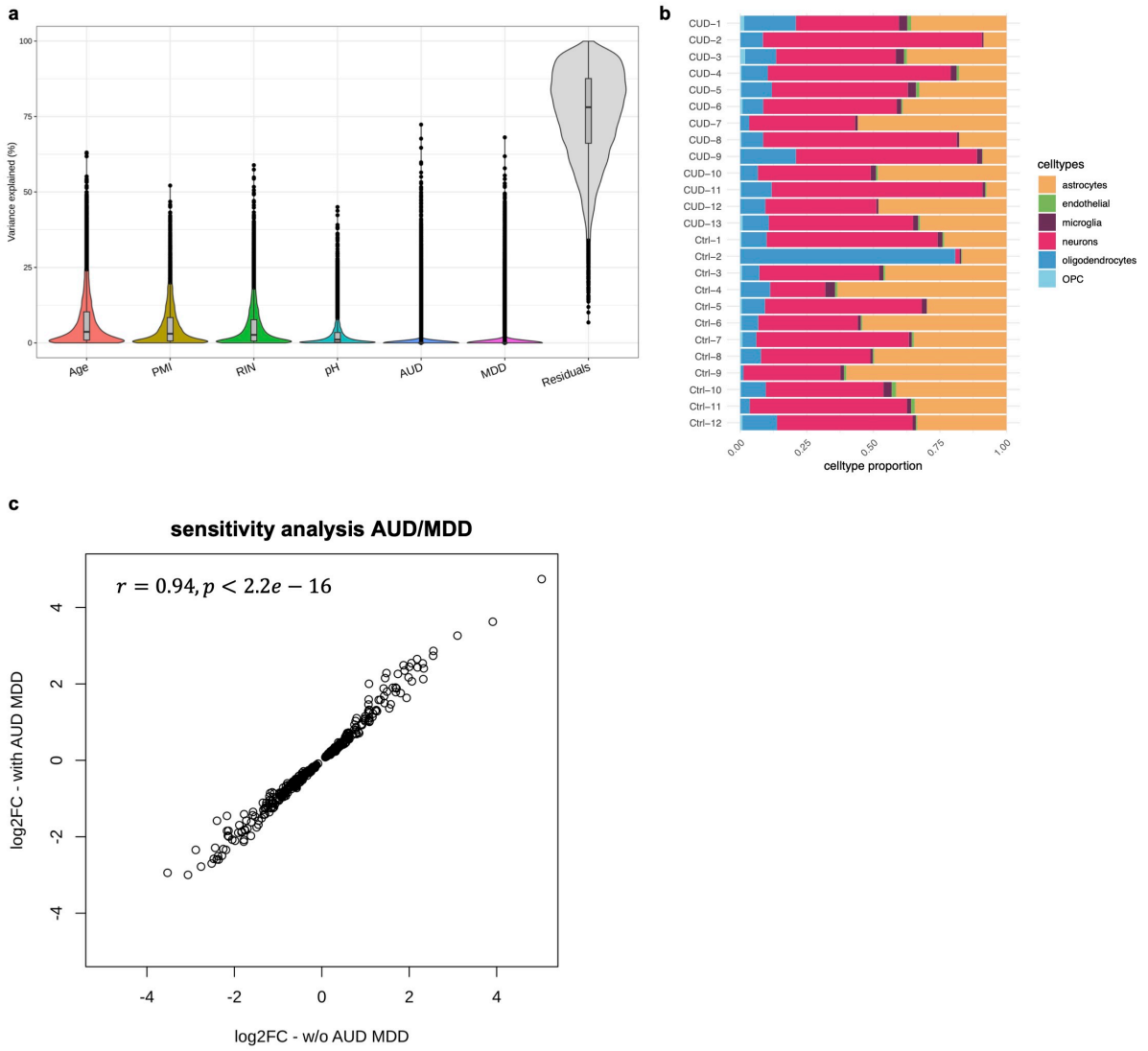
Corresponding author
Correspondence to Stephanie H. Witt.

3.11 Ethics declarations

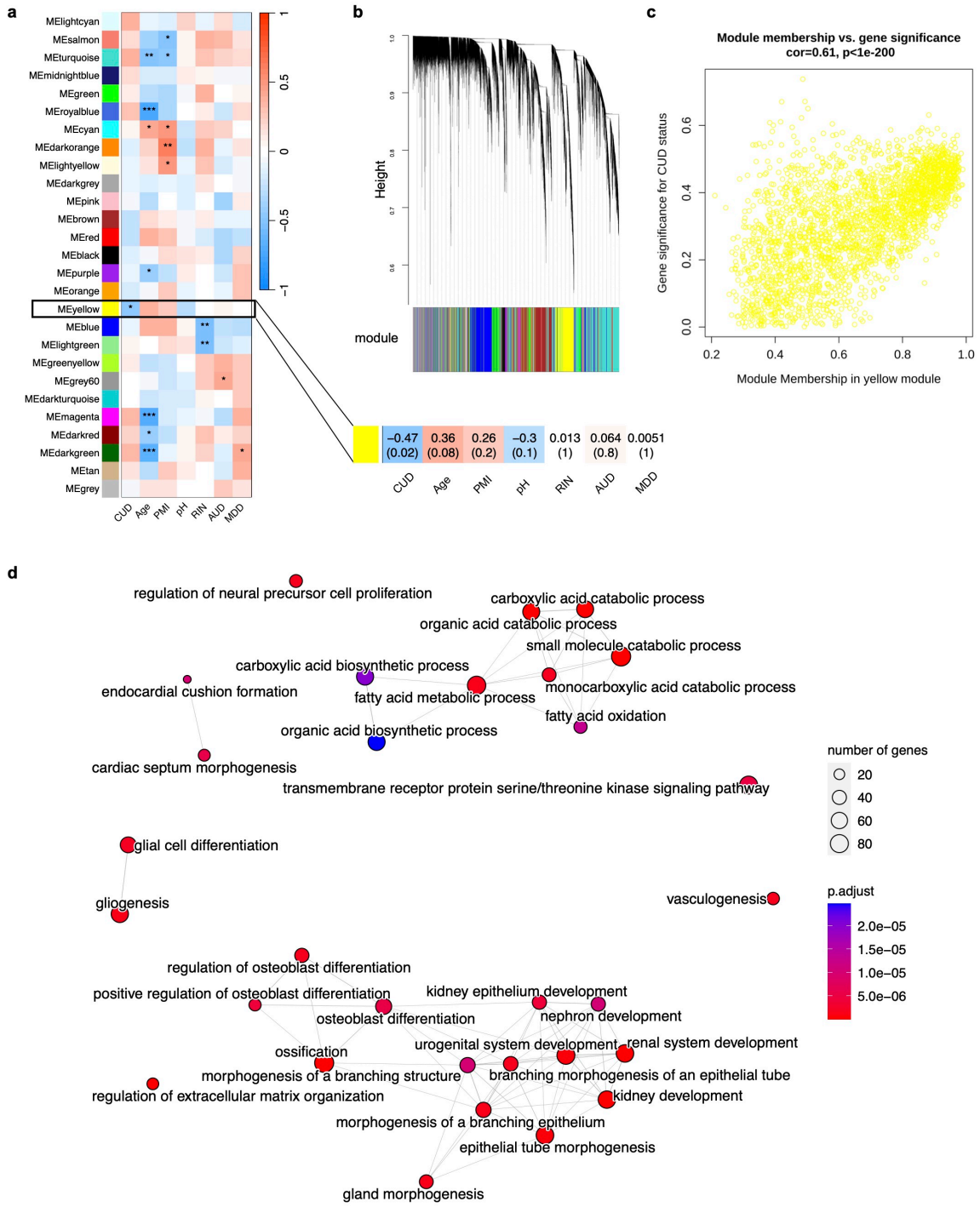
Competing interests
The authors declare no competing interests.

Ethics statement
Postmortem human brain tissue was obtained from the Douglas Bell Canada Brain Bank, where tissue sampling is performed based on their established ethical guidelines. These include informed consent from next-of-kin for tissue sampling from the deceased individual for research purposes. The experimental protocol was approved by the Ethics Committee II of the University of Heidelberg, Medical Faculty Mannheim, Germany, under the register number 2021-681. All methods were carried out in accordance with relevant guidelines and regulations.

3.12 Supplementary Figures

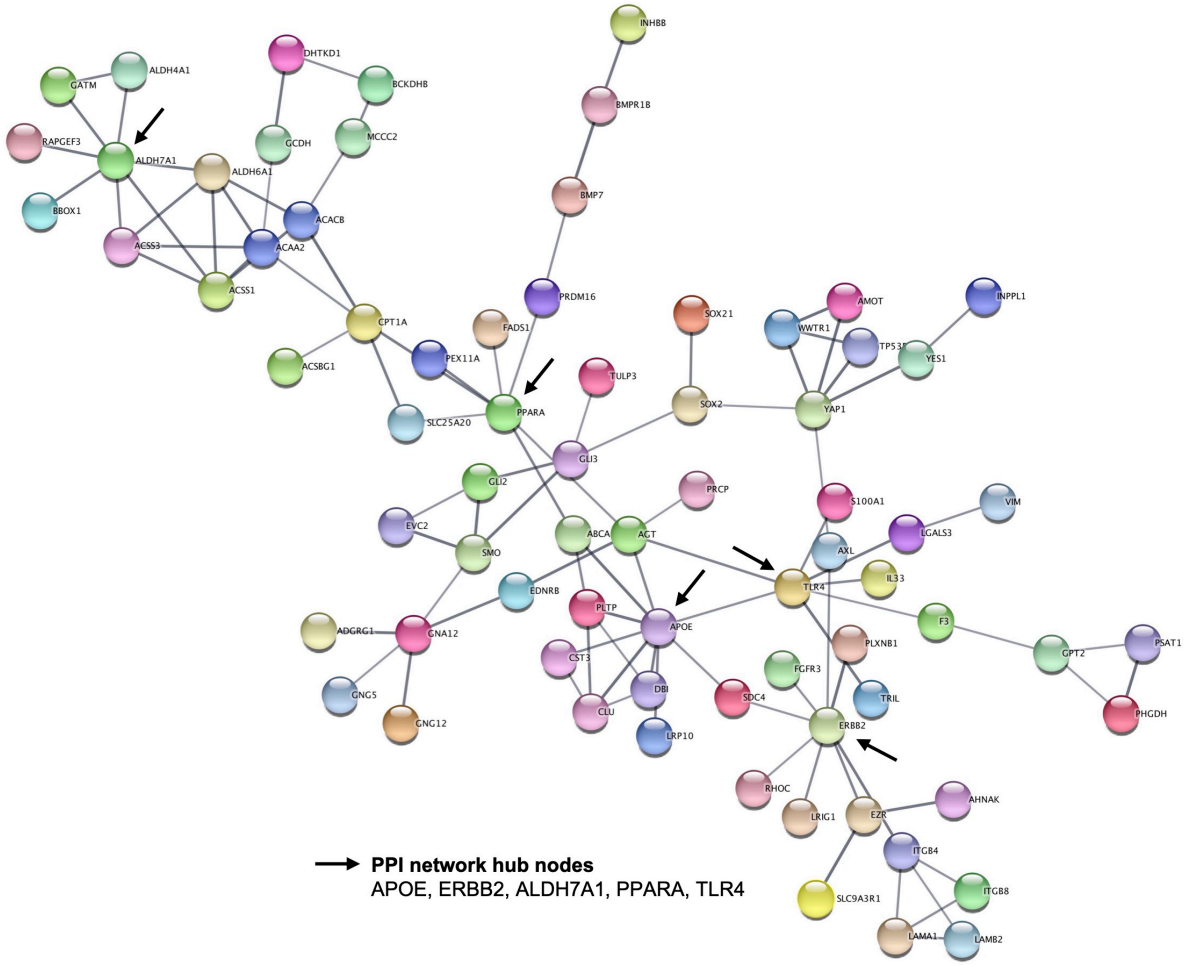


Supplementary Figure S3.1 – Differential expression analysis in Brodmann Area 9
a Variance partition analysis confirms the included covariates for the DE model in DESeq2 and reveals minimal explained variance for AUD and MDD. PMI=postmortem interval, RIN=RNA integrity number, pH=postmortem brain tissue pH value. **b** Cell type proportions in the N=25 postmortem human brain tissue samples from Brodmann Area 9 based on the estimation using CIBERSORT. CUD: N=13 cocaine use disorder samples, Ctrl: N=12 control samples. OPC=oligodendrocyte progenitor cell. **c** Correlation of log2FC estimates for nominal significant DE genes (p<0.05) in the sensitivity analysis that includes alcohol use disorder (AUD) and depressive disorder (MDD) as additional covariates into the DE model. A strong and highly significant Pearson correlation between log2FC estimates was observed (r=0.94, p<2.2e-16).

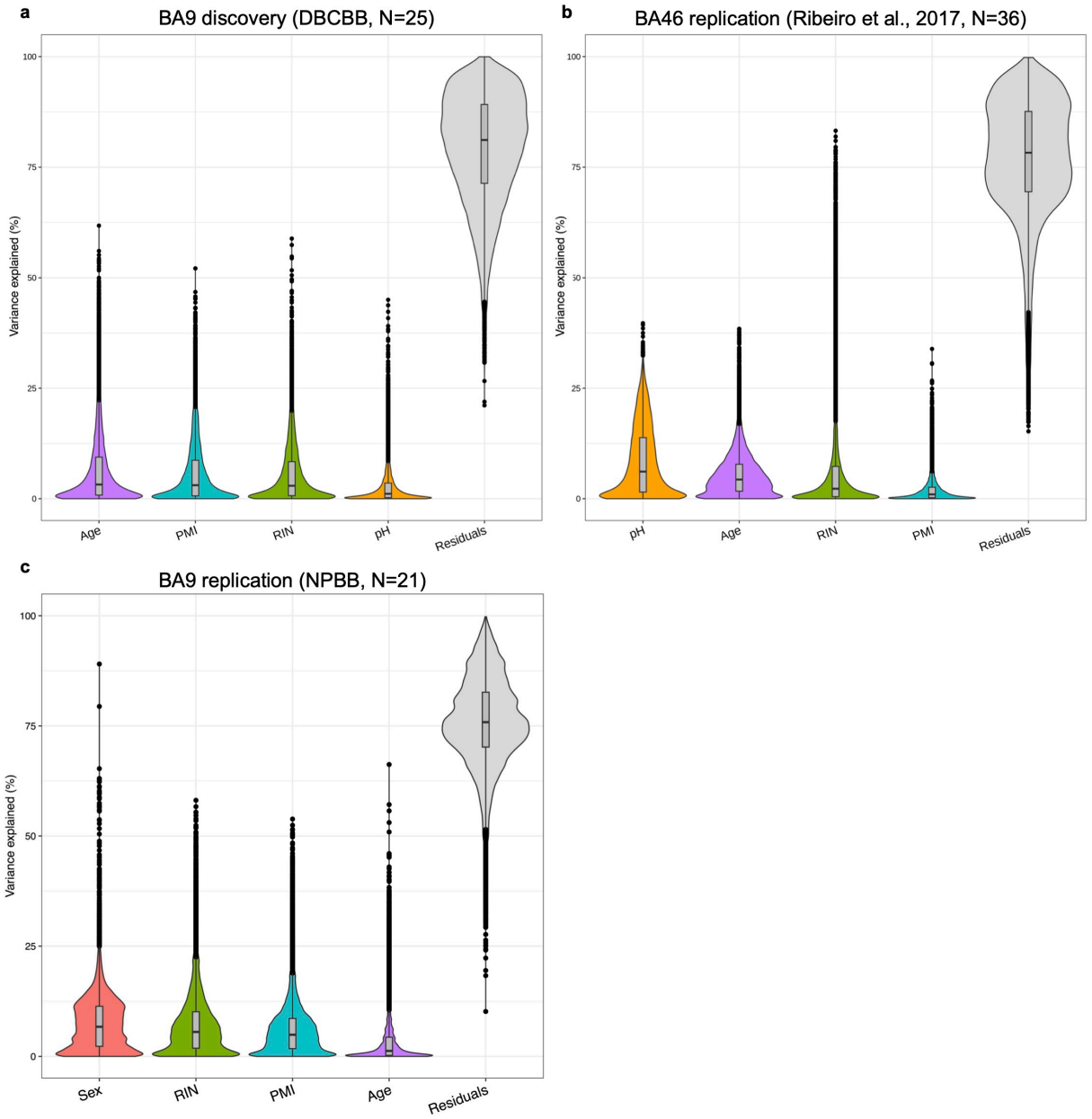


Supplementary Figure S3.2 – Weighted correlation network analysis reveals a cocaine-use disorder associated co-expression module in Brodmann Area 9
a Correlation of the N=27 identified co-expression modules from weighted correlation network analysis (WGCNA) with cocaine use disorder (CUD) status and other known phenotypes in the Brodmann Area 9 cohort. Pearson correlation of module eigengene with the phenotypic variable was color coded where red color indicates positive and blue color indicates negative correlation coefficients. Module yellow was significantly negatively correlated with CUD (r=-0.47, p=0.02). Significance of correlation is indicated using asterisks (*=p<0.05, **=p<0.01, ***=p<0.001). PMI=postmortem interval, pH=postmortem brain tissue pH value, RIN=RNA integrity number, AUD=alcohol use disorder status, MDD=depressive disorder status. **b** Dendrogram of genes and their assignment to co-expression modules in WGCNA. **c** Strong correlation of module membership and CUD gene significance for the N=2,517 genes in co-expression module yellow was identified (r=0.61, p<1e-200). **d** Results of the Gene Ontology (GO) enrichment analysis for the N=2,517

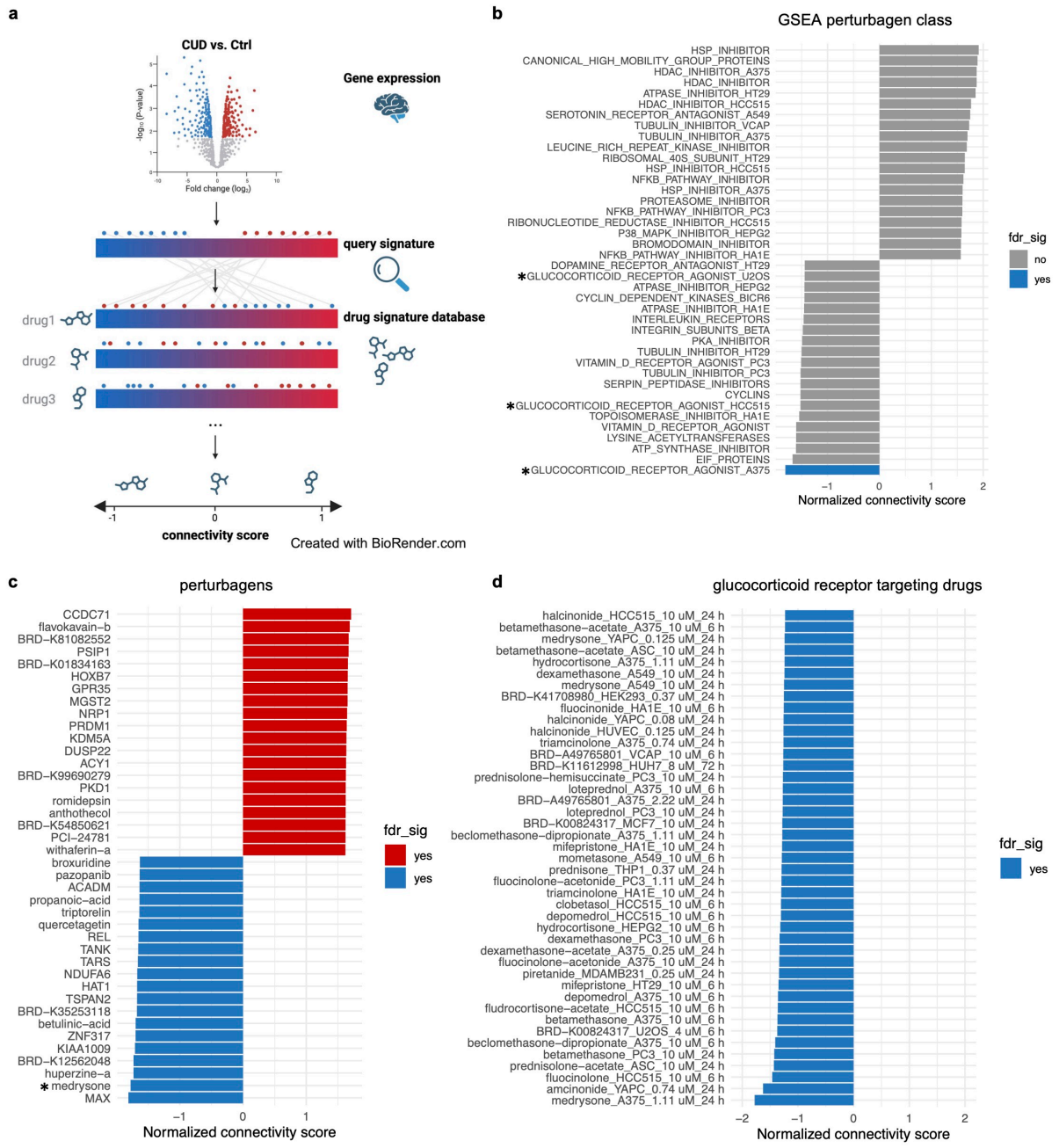
module yellow genes identified an overrepresentation in biological pathways related to development and fatty acid metabolism.



Supplementary Figure S3.3 – Protein-protein interaction network of module yellow hub genes Protein-protein interaction STRING network of the top 10% module yellow hub genes reveals several highly connected network hub nodes such as APOE, ERBB2, and ALDH7A1, PPARA, and TLR4. Network hub nodes are highlighted by arrows.

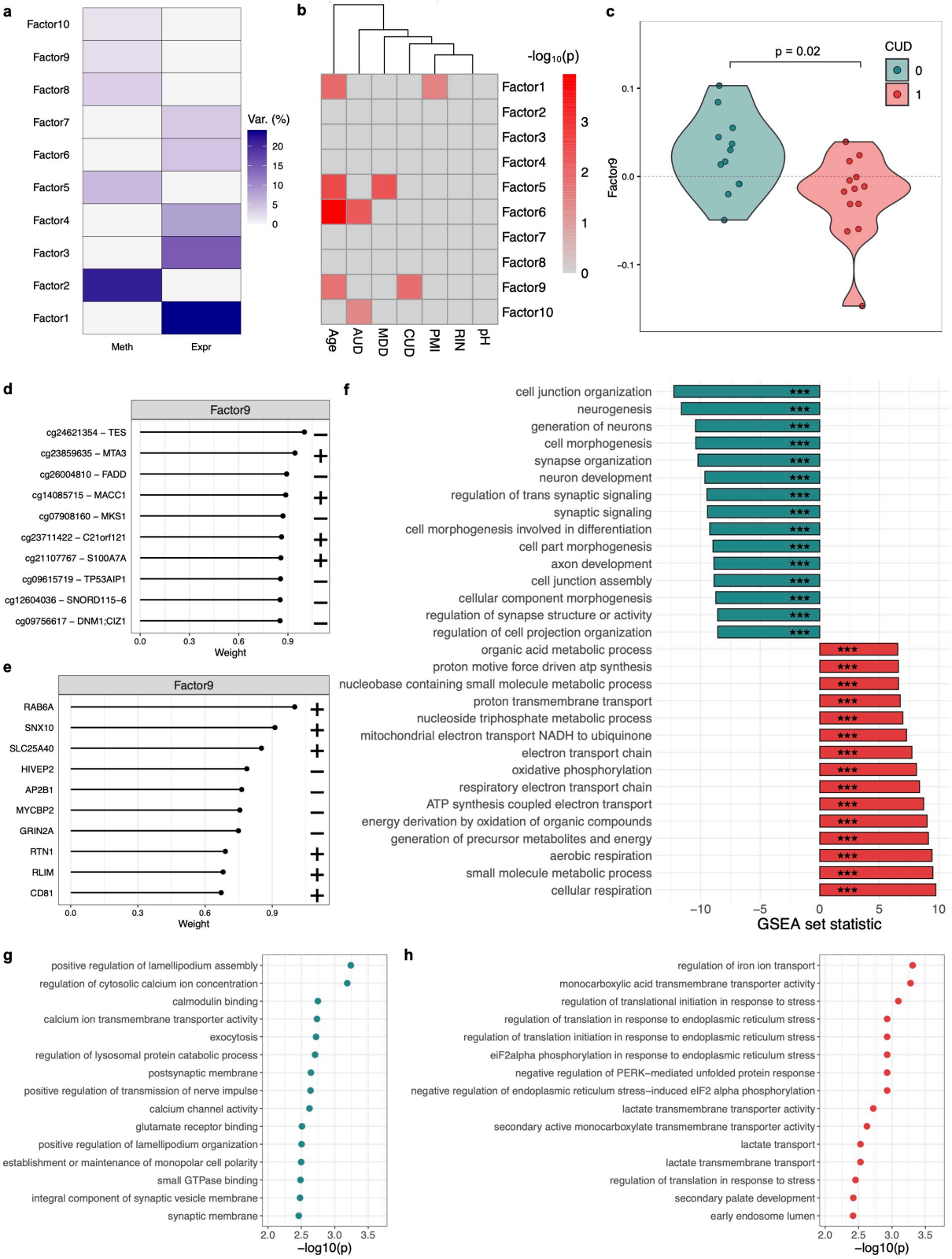


Supplementary Figure S3.4 – Variance explanation for DESeq2 model covariates across discovery and replication datasets. Variance partition analysis results for variables included as covariates into the DESeq2 differential gene expression analysis of **a** DBCBB RNA-seq discovery dataset from Brodmann Area 9 (BA9, N=25) **b** neuron-specific RNA-seq replication dataset from Brodmann Area 46 (N=36) and **c** BA9 RNA-seq replication dataset from NPBB (N=21).



Supplementary Figure S3.5 – Drug repositioning analysis based on the differential gene expression signature in cocaine use disorder suggests glucocorticoid receptor targeting drugs as a potential pharmacotherapy

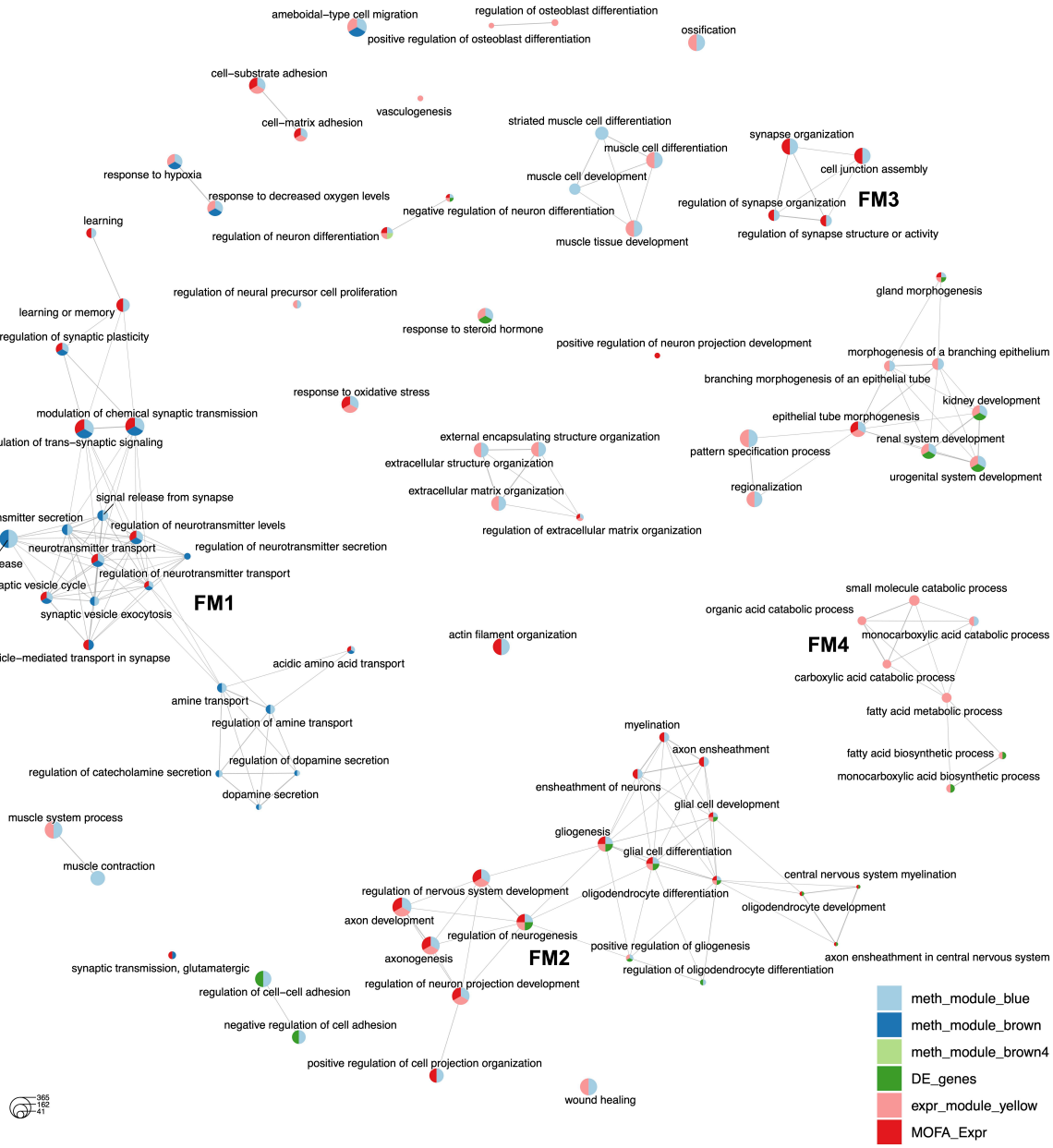
a Schematic overview of the drug repositioning analysis approach based on Connectivity Map (CMap) using the LINCS L1000 perturbagen reference database. The top N=150 upregulated and N=150 downregulated differentially expressed genes in cocaine use disorder were used as the query signature. **b** Results of the GSEA perturbagen class analysis from CMap. The top N=20 findings with positive and negative normalized connectivity score (NCS) are shown. Asterisks highlight findings related to glucocorticoid receptor agonists in different cell lines. **c** CMap results for individual perturbagen signatures that are most similar or dissimilar to the query signature. The top N=20 perturbagens with positive and negative NCS are shown. Asterisk highlights the glucocorticoid receptor targeting drug medrysone. **d** NCS for all glucocorticoid receptor targeting perturbagen drugs available in CMap were extracted indicating consistently significant negative NCS across cell lines compared to the query signature from Brodmann Area 9. Blue color indicates significance (fdr_sig, q<0.05) after FDR-correction for multiple testing.



Supplementary Figure S3.6 – Integrated analysis of DNA methylation and gene expression data from Brodmann Area 9 using multi-omics factor analysis provides further evidence for synaptic signaling and metabolic alterations in cocaine use disorder

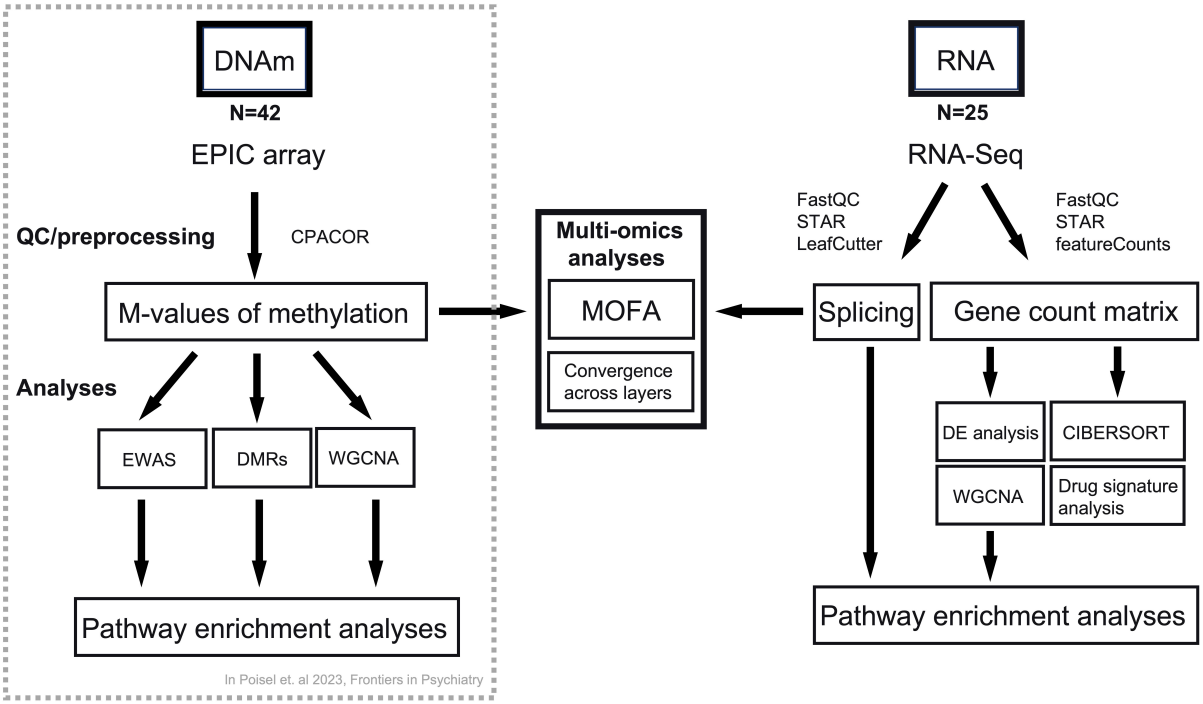
a Variance explanation (in %) in the DNA methylation (Meth, N=20,000 most variable promoter CpG sites) and expression dataset (Expr, N=20,000 most variable genes) for each of the identified factors from multi-omics factor analysis (MOFA). **b** Association of MOFA factors with cocaine use disorder (CUD) and known covariates in the postmortem human brain cohort. AUD = alcohol use disorder, MDD/NOS = major depressive disorder/depressive disorder not otherwise specified, PMI=

postmortem interval, RIN = RNA integrity number, pH = brain tissue pH value. Red color indicates statistically significant association with the phenotype (p<0.05). **c** Significant difference in factor 9 values was observed between CUD cases and Ctrl (p=0.02). Top 10 **d** CpG sites (Meth) and **e** genes (Expr) having the strongest absolute weights on factor 9. +: positive weight, -: negative weight. **f** Results of the gene set enrichment analysis (GSEA) on the expression weights on factor 9 separated by positive and negative weights. Significance estimates inside bars are FDR-adjusted q-values (***, q<0.001). For the DNA methylation dataset, results of the GO enrichment analysis for CpG sites with strongest **g** negative weights and **h** positive weights on factor 9 are shown.



Supplementary Figure S3.7 – Integrative Gene Ontology enrichment analysis based on DNA methylation and expression signatures in Brodmann Area 9 reveals multiple functional modules associated with cocaine use disorder
Enrichment map of the Gene Ontology (GO) enrichment analysis using cocaine use disorder (CUD) associated gene lists derived from individual and integrative analyses of DNA methylation and gene expression signatures in Brodmann Area 9. Convergent evidence at the biological pathway

level leads to the emergence of functional modules (FM) with interconnected GO terms where statistically significant (q<0.05) enrichment was detected for multiple gene lists.



Supplementary Figure S3.8 – Analysis workflow of the multi-omics analysis of DNA methylation (DNAm) and gene expression (RNA) in Brodmann Area 9 in cocaine use disorder
Analysis of the DNA methylation (DNAm) dataset from the N=42 Brodmann Area 9 samples has been performed in Poisel et al., 2023. Transcriptome-wide gene expression analysis from RNA-sequencing in a subset of N=25 individuals of this cohort and multi-omics data integration is performed in the present study. CPACOR=Control Probe Adjustment and reduction of global CORrelation (Lehne et al., 2015), EWAS=epigenome-wide association study, DMRs=differentially methylated regions, WGCNA=weighted gene co-expression network analysis, MOFA=multi omics factor analysis, DE=differential expression. The DNAm analysis was previously published in Poisel et al., Frontiers in Psychiatry, 2023.

3.13 Supplementary Tables

Due to genome-wide table sizes, for full supplementary tables, please refer to the online version of the supplementary information:

https://static-content.springer.com/esm/art%3A10.1038%2Fs41398-024-03139-9/MediaObjects/41398_2024_3139_MOESM2_ESM.xlsx

Supplementary Table S3.1. Comparative overview on demographic data for the subset of N=25 individuals with RNA-seq data and the full cohort of N=42 individuals with DNA methylation data in Brodmann Area 9. Tissue was obtained from the Douglas Bell Canada Brain Bank (DBCBB), Montreal, Canada.

Supplementary Table S3.2a. Inferred cell-type fractions from the cell-type deconvolution analysis in CIBERSORT.

Supplementary Table S3.2b. Results of the Bayesian estimation for the difference in group means (CUD vs. Ctrl) using the cell-type distributions from CIBERSORT. The R function bayes.t.test from the BayesianFirstAid package was used to estimate the 95% high density interval.

Supplementary Table S3.3a. Summary Statistics of the differential expression analysis for cocaine use disorder (CUD) in BA9. Column names correspond to default labels from DESeq2 output.

Supplementary Table S3.3b. Deviance statistics of the differential expression analysis for cocaine use disorder (CUD) in BA9. Deviance was extracted from DESeq2 model statstics, "no" indicates full model (~ CUD + Age + RIN + pH + PMI) minus respective variable. R2 was calculated as 1-residual deviance.

Supplementary Table S3.4. Positive enrichment results of the gene-set enrichment analysis (fgsea) using the differential expression results stat-column.

Supplementary Table S3.5. Negative enrichment results of the gene-set enrichment analysis (fgsea) using the differential expression results stat-column.

Supplementary Table S3.6. Gene statistics for the CUD-associated WGCNA module yellow and module hub genes.

Supplementary Table S3.7. GO enrichment analysis for genes in WGCNA module yellow.

Supplementary Table S3.8. Results of the differential splicing analysis using LeafCutter - clusters with significant differential splicing in CUD.

Supplementary Table S3.9. Results of the differential splicing analysis using LeafCutter - dPSI changes of individual introns.

Supplementary Table S3.10. GO enrichment analysis using genes harboring differentially spliced clusters (5% FDR for significance of differential splicing).

Supplementary Table S3.11. Input gene list for the drug signature analysis using CMap (https://clue.io/).

Supplementary Table S3.12. Correlation of the subset of genes for which both differential expression ($p<0.05$) and differential methylation ($p<0.01$) was detected.

Supplementary Table S3.13. Feature weights for DNA methylation and gene expression features on CUD-associated MOFA factor 9.

Supplementary Table S3.14. Results of the gene set enrichment analysis on positive and negative expression feature weights on CUD-associated MOFA factor 9.

Supplementary Table S3.15. Results of the GO enrichment analysis on positive and negative methylation feature weights on CUD-associated MOFA factor 9.

Supplementary Table S3.16. Gene sets for the analysis of converging evidence at the pathway level.

Supplementary Table S3.17. Demographic information for the BA9 replication dataset (N=21) based on postmortem brain tissue from the National PTSD Brain Bank (NPBB).

4 STUDY 3 - A MULTI-OMICS AND CELL TYPE-SPECIFIC CHARACTERIZATION OF THE VENTRAL STRIATUM IN HUMAN COCAINE USE DISORDER[3]

4.1 Highlights

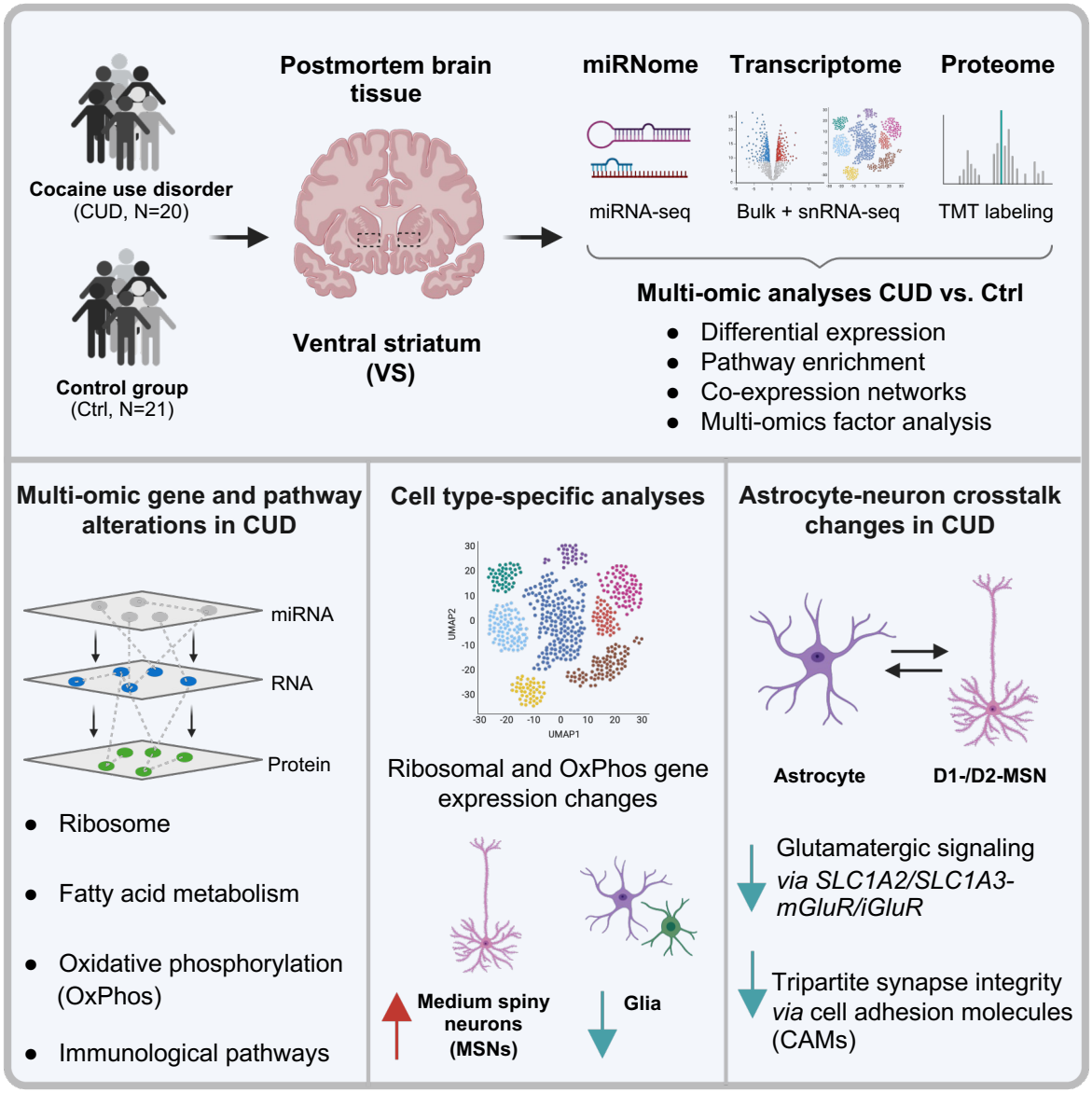
- Integrative multi-omics analysis of miRNA, RNA, and protein changes in human CUD
- Cell type specificity of transcriptional changes in CUD identified by snRNA-seq
- Strong deregulation patterns in ventral striatal astrocytes and medium spiny neurons
- Altered astrocyte-neuron crosstalk implying glutamatergic and cell-cell adhesion changes

4.2 Summary

Epigenome, transcriptome, and proteome analyses of postmortem brains have revealed initial molecular insights into cocaine use disorder (CUD). However, the inter-relationship between these omics and the contribution of individual cell types remains largely unknown. We present an in-depth analysis of molecular changes in the ventral striatum in CUD at multi-omics and single-cell resolution. Integrative multi-omics analyses of microRNA sequencing (microRNA-seq), RNA sequencing (RNA-seq), and proteomics datasets in 41 individuals and single-nuclei RNA-seq in a subset of 16 individuals revealed conserved deregulation of metabolic pathways, oxidative phosphorylation, and glutamatergic signaling. Cell type-specific analyses identified inverse metabolic pathway deregulation patterns in glial and neuronal cells, notably in astrocytes and medium-spiny neurons (MSNs). Characterizing astrocyte-neuron crosstalk revealed altered glutamatergic and cell-cell adhesion signaling in CUD. By applying a comprehensive multi-omics analytical framework, our study provides novel insights into CUD-associated molecular changes in the ventral striatum highlighting the perturbation of astrocytes, MSNs, and their crosstalk in CUD.

[3] Publication: Zillich, E., Artioli, A., Rossetti, A. C., Avetyan, D., Belschner, H., Frank, J., Stein, F., Schwarz, J. J., Mechawar, N., Turecki, G., Nöthen, M. M., Hansson, A. C., Witt, C. C., Rietschel, M., Koch, P., Spanagel, R., Zillich, L., & Witt, S. H. (2025). A multi-omics and cell type-specific characterization of the ventral striatum in human cocaine use disorder. *Cell Reports*, *44*(2). https://doi.org/10.1016/j.celrep.2025.115332

4.3 Graphical abstract



4.4 Keywords

cocaine use disorder, multi-omics, addiction, postmortem brain tissue, proteomics, single-nuclei RNA-seq, miRNA-seq

4.5 Research Topic(s)

CP: Neuroscience

4.6 Introduction

Individuals with cocaine use disorder (CUD) present with an excessive intake of the psychostimulant cocaine despite negative consequences, strong cocaine

craving, and relapse after periods of abstinence (American Psychiatric Association, 2013). In 2021, a total of 21.6 million individuals used cocaine worldwide (UNODC, 2023b), with around 20% of cocaine users transitioning from episodic cocaine use to use disorder during their lifetimes (Lopez-Quintero et al., 2011). Current treatment options are limited and retrospective analyses suggest that less than 25% of treatment-seeking patients remain abstinent after completing an inpatient treatment program (Paliwal et al., 2008). A deeper understanding of the neurobiological mechanisms of CUD is essential to provide the basis for the development of mechanism-based interventions.

The brain is assumed to be the most prominently affected organ in the development and maintenance of CUD. Neuroimaging studies revealed structural brain alterations such as reduced gray matter volume in CUD, and functional changes in neurocircuit connectivity between different brain regions of the reward system have been described (Ceceli et al., 2023; Connolly et al., 2013; Ersche et al., 2011; Gaudreault et al., 2023; Goldstein & Volkow, 2002; Yuzheng Hu et al., 2015; King et al., 2022). Cocaine-induced epigenetic and transcriptional changes in the human brain were proposed as molecular mechanisms involved in the formation of structural and functional neurocircuit changes in individuals with CUD (Robison & Nestler, 2011). Investigating molecular signatures of CUD in the human brain thus depicts an important approach toward a better understanding of the underlying disease processes (Fernàndez-Castillo et al., 2022). Previous studies in postmortem human brain tissue have reported on molecular alterations of the epigenome, transcriptome, and proteome in CUD (Fernàndez-Castillo et al., 2022; Mews et al., 2023; Poisel et al., 2023; Ribeiro et al., 2017; Tondo et al., 2021; Vaillancourt, Chen, et al., 2021; Vaillancourt, Yang, et al., 2021; Viola et al., 2019; Zhou et al., 2011). Epigenome-wide studies have so far mainly focused on DNA methylation and showed CUD-associated differential methylation in multiple addiction-relevant brain regions such as the prefrontal cortex (PFC) (Poisel et al., 2023), the ventral striatum (VS) (Vaillancourt, Chen, et al., 2021), and the caudate nucleus (CN) (Vaillancourt, Yang, et al., 2021). Differential methylation levels were detected in genes involved in dopamine metabolism, such as tyrosine hydroxylase (Vaillancourt, Chen, et al., 2021), while, at the pathway level, epigenetic changes were related to transcription factor activity and synaptic signaling (Poisel et al., 2023; Vaillancourt, Chen, et al., 2021). Another domain of epigenetic regulators are micro-RNAs (miRNAs), small RNA molecules that bind to complementary nucleotide sequences on mRNAs, thereby regulating mRNA degradation and translation rate to proteins (Jonas & Izaurralde, 2015). While associations with CUD were identified in peripheral blood for miRNAs such as miR-124 and miR-184 (Viola et al., 2019), differential miRNA expression remains understudied in the human CUD brain. At the transcriptomic scale, multiple studies have identified CUD-associated changes in RNA levels in cortical (Ribeiro et al., 2017), limbic (Zhou et al., 2011), and striatal brain regions (Mews et al.,

2023). The most prominent findings include alterations of transcripts and co-expression networks involved in neuroplasticity, neuroinflammation, and mitochondrial respiration (Mews et al., 2023; Ribeiro et al., 2017; Zhou et al., 2011). Further, investigating the proteome is particularly important for evaluating altered neurobiological functions in CUD as protein-protein interactions depict a key component of cellular signaling. Proteomic analysis of the human prefrontal cortex in CUD revealed differential expression of proteins involved in neuroinflammation and myelination, supporting the neuroimaging findings of white matter deficits in postmortem brain (Tondo et al., 2021).

While single-omics studies are valuable in characterizing disease-associated changes in a class of biological molecules such as RNAs or proteins, it remains unclear to what extent these molecular alterations are conserved across layers of biological regulation. Integrative multi-omics analysis of epigenomic regulation, the transcriptome, and the proteome addresses the inter-regulated nature of biological processes, thereby depicting a powerful tool to uncover molecular mechanisms of biological deregulation. Further, cell type-specific associations cannot be sufficiently deduced in bulk-level omics-wide association studies, increasing the need for analyses at single-cell resolution. Initial single-nuclei RNA sequencing (snRNA-seq) studies have been performed in different brain regions, mainly focusing on rodent models of CUD (Phillips et al., 2023; Savell et al., 2020; Zhou et al., 2023; Zillich et al., 2025). While previous findings from bulk transcriptomic studies such as differential expression of neuroplasticity genes were confirmed in snRNA-seq approaches, results were strongly cell ype dependent, highlighting the importance of further analyses at single-cell resolution.

In the present study, we addressed the limitations of single-omics association studies by performing an integrative multi-omics analysis of miRNA sequencing (miRNA-seq), RNA sequencing (RNA-seq), and proteomic data from the same postmortem brain tissue cohort followed by a cell type-specific investigation of transcriptomic signatures. Analyses were performed in a collection of $n = 41$ postmortem human brain samples of the VS, an important brain region of the neurocircuitry of addiction involved in reward and reinforcement processing (Cox & Witten, 2019). We performed bulk-level multi-omics analyses with two main objectives: the identification of CUD associations at multiple individual molecular levels and the investigation of the inter-relationship of findings across layers of biological regulation. We additionally performed snRNA-seq to evaluate the cell type specificity of transcriptomic changes in CUD. We finally integrated our human snRNA-seq CUD dataset with rodent snRNA-seq data to identify potential converging evidence between human CUD and a controlled cocaine-exposure study in rats. Here, we present an in-depth characterization of CUD in postmortem

human brain at both bulk and single-nuclei resolution identifying metabolic, synaptic, and immunological changes as molecular hallmarks of the VS in CUD.

4.7 Results

We generated miRNA-seq (n = 40), RNA-seq (n = 40), proteomics (n = 40), and snRNA-seq data (n = 16) in a collection of n = 41 postmortem human brain tissue samples (n = 20 individuals with CUD, n = 21 without CUD). A graphical summary of the study design and analyses is provided in Figures 4.1A and 4.1B. Dataset availability for each of the VS samples is shown in Figure S4.1A. To assess potential systematic differences in phenotypes affecting miRNA, RNA, or protein levels, we evaluated donor demographics. Besides the cause of death, with CUD cases having a significantly higher rate of suicide (p = 0.002, Table S4.1), no significant phenotype differences were found between CUD cases and controls (Ctrls). After principal component analysis (PCA)-based quality control, 38 individuals remained for miRNA-seq, RNA-seq, and proteomic data analyses (Figure S4.1B). To address potential confounding from differential cell type fractions, we performed cell type deconvolution using CIBERSORT in the bulk RNA-seq dataset (Figures S4.1C and S4.1D; STAR Methods). This confirmed no significant differences in cell type proportions between the 17 CUD cases and 21 Ctrls (Table S4.2A).

4.7.1 Individual analysis of bulk-level miRNA-seq, RNA-seq, and proteomic data in the VS suggests ribosomal and immunological alterations in CUD

We performed individual-level differential expression analyses of miRNA-seq, RNA-seq, and proteomics datasets (n = 38 individuals) to identify associations with CUD for each of the three molecular levels. Analysis of miRNA-seq data revealed 16 upregulated and 21 downregulated miRNAs in CUD at nominal significance ($p < 0.05$), but no miRNA association remained significant after multiple testing correction. The strongest association for upregulated miRNAs was found for hsa-miR-6833-3p (log2 fold change [FC] = 0.64, p = 1.67e−02, q = 1), while hsa-miR-508-5p (log2FC = −2.13, p = 2.18e−03, q = 1) was the top finding among downregulated miRNAs in CUD (Figure 4.1C; Table S4.2B). The transcriptome (RNA-seq) analysis identified 36 transcriptome-wide significant differentially expressed genes (DEGs, 5% false discovery rate [FDR]; Figure 4.1D; Table S4.2C). Cytokine signaling regulator *SOCS3* (log2FC = 2.47, p = 9.63e−07, q = 7.38e−03) and sialyltransferase *ST6GAL2* (log2FC = −0.30, p = 2.90e−07, q = 5.71e−03) were the top findings of up- and downregulated DEGs, respectively.

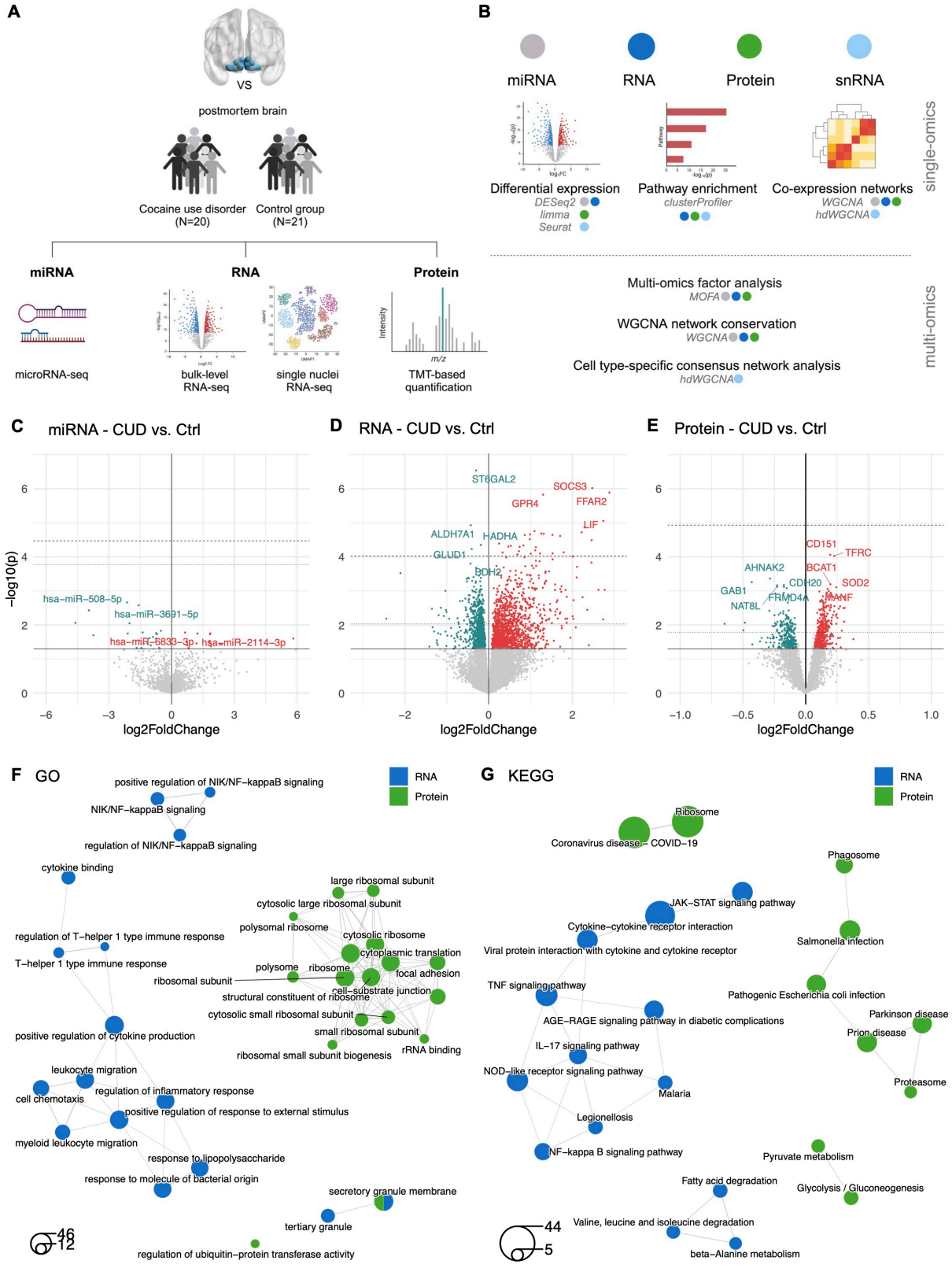


Figure 4.1 - Multi-omics characterization of miRNA, RNA, and protein changes in the VS in CUD

(A) Graphical summary of the study. (B) Summary of analysis methods. (C–E) Bulk-level analysis results of (C) miRNA, (D) RNA, and (E) protein differential expression (DE) analyses. Dashed line indicates 5%-FDR and dotted line indicates 25%-FDR significance levels; |log2FC| >0.07 (5% change in expression levels). (F) Semantic clustering of statistically significant (FDR q < 0.05)

results from GO enrichment analysis at the RNA (blue) and protein levels (green). (G) Results of the same analysis approach for KEGG pathways.

Using tandem mass tag (TMT)-based proteomics, we detected 765 differentially expressed proteins (DEPs) in CUD with surface glycoprotein CD151 (log2FC = 0.19, p = 8.68e−05, q = 0.20) and AHNAK2 involved in calcium signaling (log2FC = −0.28, p = 4.33e−04, q = 0.20) as top up- and downregulated DEPs, respectively (Figure 4.1E; Table S4.2D). Of the 765 DEPs, a total of 469 were up- and 296 were downregulated at nominal significance ($p < 0.05$), while no DEP remained statistically significant after multiple testing correction.

To characterize altered biological pathways in the VS in a conservative but sufficiently sized gene set, we focused on DEGs and DEPs passing a relatively lenient but still statistically stringent significance threshold of 25% FDR (q < 0.25). In the set of 718 DEGs and 282 DEPs associated with CUD at q < 0.25, we performed pathway enrichment analyses using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases as reference (Table S4.2E). Clustering of significant GO terms based on biological similarity revealed a large functional GO term module specific to the RNA level related to immune signaling (Figure 4.1F). Protein-specific enrichment was found among ribosomal pathways that formed another highly connected GO term module. KEGG pathway analysis confirmed the overrepresentation of DEGs and DEPs within ribosomal and immune processes while suggesting additional metabolic changes related to fatty and amino acid metabolism (Figure 4.1G). Further, KEGG analysis revealed a DEP-specific module associated with neurodegenerative diseases that show overlapping symptoms with CUD, such as brain atrophy (Ersche, Jones, et al., 2013). In an overlap analysis of pathway associations in the transcriptomic and proteomic dataset at 25% FDR, the GO terms “secretory granule membrane” and “ficolin-1-rich granule” emerged as a significant finding both at the RNA and protein level (Figure 4.1F), while, in a more lenient analysis using nominally significant DEGs and DEPs, a strong overlap at fatty acid metabolism pathways was additionally observed (Table S4.2E).

To further elaborate on the relationship between transcriptomic and proteomic profiles of the VS, we performed a transcriptome-proteome correlation analysis using expression information from 3,935 genes for which both RNA and protein data were available (see STAR Methods). Based on mean RNA and protein expression levels across samples, we observed a moderate Pearson correlation of r = 0.43 ($p < 2.2e{-}16$) between transcriptome and proteome (Figures S4.2A–S4.2C; Table S4.3). In addition, we found the overall correlation to be independent of CUD status with correlation coefficients of r = 0.43 in both CUD and Ctrl conditions (Figure S4.2D). To identify genes for which RNA levels correlate with protein levels particularly well or poorly, we selected the genes with the strongest

positive and negative correlation coefficients of RNA and protein levels in the VS. Here, strong concordance of RNA and protein levels was observed for synaptic signaling genes, while an inverse relationship between RNA and protein levels was most prominent among oxidative phosphorylation genes (Figures S4.2E and S4.2F). Next, we aimed to provide further insights into synchronization and desynchronization patterns of RNA-protein correlation by assessing the difference between gene-level transcriptome-proteome correlation coefficients calculated in CUD and Ctrl samples individually. Among genes showing the strongest positive differences in RNA-protein correlation coefficients; i.e., genes with synchrony between RNA and protein expression levels in CUD, we found several synaptic genes such as SV2C ($\Delta R = 0.81$), CAMK2D ($\Delta R = 0.78$), and STXBP3 ($\Delta R = 0.75$; Table S4.3). In contrast, among genes with reduced synchrony between RNA and protein expression in CUD, we observed oxidative phosphorylation genes including NDUFS6 ($\Delta R = -1.21$), NDUFV2 ($\Delta R = -0.82$), and NDUFB4 ($\Delta R = -0.77$).

4.7.2 Proteo-transcriptomic co-expression network analysis reveals conservation of CUD-associated metabolic gene networks across RNA and protein levels

Profiling gene networks allows the identification of co-regulated gene expression programs that often provide more information about altered biological processes than deregulation patterns of individual genes. To construct co-expression modules at the multi-omics scale, we first performed weighted correlation network analyses (WGCNAs) in miRNA, mRNA, and protein datasets individually and then performed integrative analysis of networks across omics. In the miRNA expression dataset, we found eight co-expression modules for which no significant association with CUD was detected. Network construction in RNA-seq data revealed 54 co-expression modules of which four ("grey60," "lightcyan1," "skyblue3," and "midnightblue") were significantly associated with CUD and also showed a significant association with CUD status when adjustment for covariates was performed in a linear regression model (Figure 4.2A; Tables S4.4A and S4.4B). RNA module grey60 ($r = -0.45$, $p = 0.004$) was enriched for astrocytic and fatty acid metabolism pathways and a significant positive association with astrocytes was found ($r = 0.46$, $p = 0.004$). Module lightcyan1 ($r = 0.34$, $p = 0.04$) consisted of genes overrepresented in chromatin remodeling pathways, while skyblue3 was associated with aldehyde dehydrogenase and oxidoreductase activity (Table S4.4C; Figure 4.2C). In the proteomic dataset, we found 23 protein co-expression modules, of which two displayed a significant correlation and covariate-adjusted association with CUD including modules "yellow" and "brown" (Figure 4.2B; Tables S4.4A and S4.4B). Protein modules yellow and brown displayed significant enrichment for pathways previously identified at the RNA level, including fatty acid metabolism (yellow + brown) and oxidoreductase activity

(brown; Figure 4.2C; Table S4.4C). Biological functions unique to the protein level were synaptic vesicle (yellow) and ribosomal (brown) pathways. Finally, we were interested in whether the identified WGCNA modules were enriched for DEGs and DEPs (both $q < 0.25$) suggesting their dynamic change in CUD. We found statistically significant enrichment of DEGs in all four CUD-associated RNA modules most prominently among downregulated DEGs (Figure 4.2D). Further, significant enrichment of DEPs was found in the two protein modules indicating differential network activity of RNA and protein co-expression modules in CUD (Figure 4.2E).

4.7.3 Factor-analysis-based multi-omics integration of miRNA-seq, RNA-seq, and proteomic datasets confirms metabolic changes as a key hallmark of the CUD brain

As an additional multi-omics integration analysis of miRNA-seq, RNA-seq, and proteomic datasets, we performed multi-omics factor analysis (MOFA). Here, we aimed to identify a latent factor representation of our high-dimensional CUD dataset by an integrative "in-parallel" analysis of the three omics datasets. MOFA inferred 10 latent factors (Figure 4.3A). In total, the MOFA model explained 34%, 71%, and 61% of the variance in the miRNA-seq, RNA-seq, and proteomic datasets, respectively. Correlation of factors with covariates revealed a significant association of factor 10 with CUD status ($r = 0.33$, $p = 0.04$; Figures 4.3B and 4.3C), a factor enriched for pathways involved in synaptic signaling and oxidative phosphorylation both having been reported as deregulated biological processes in human CUD (Ribeiro et al., 2017; Zhou et al., 2011; Zillich et al., 2024; Zillich et al., 2025).

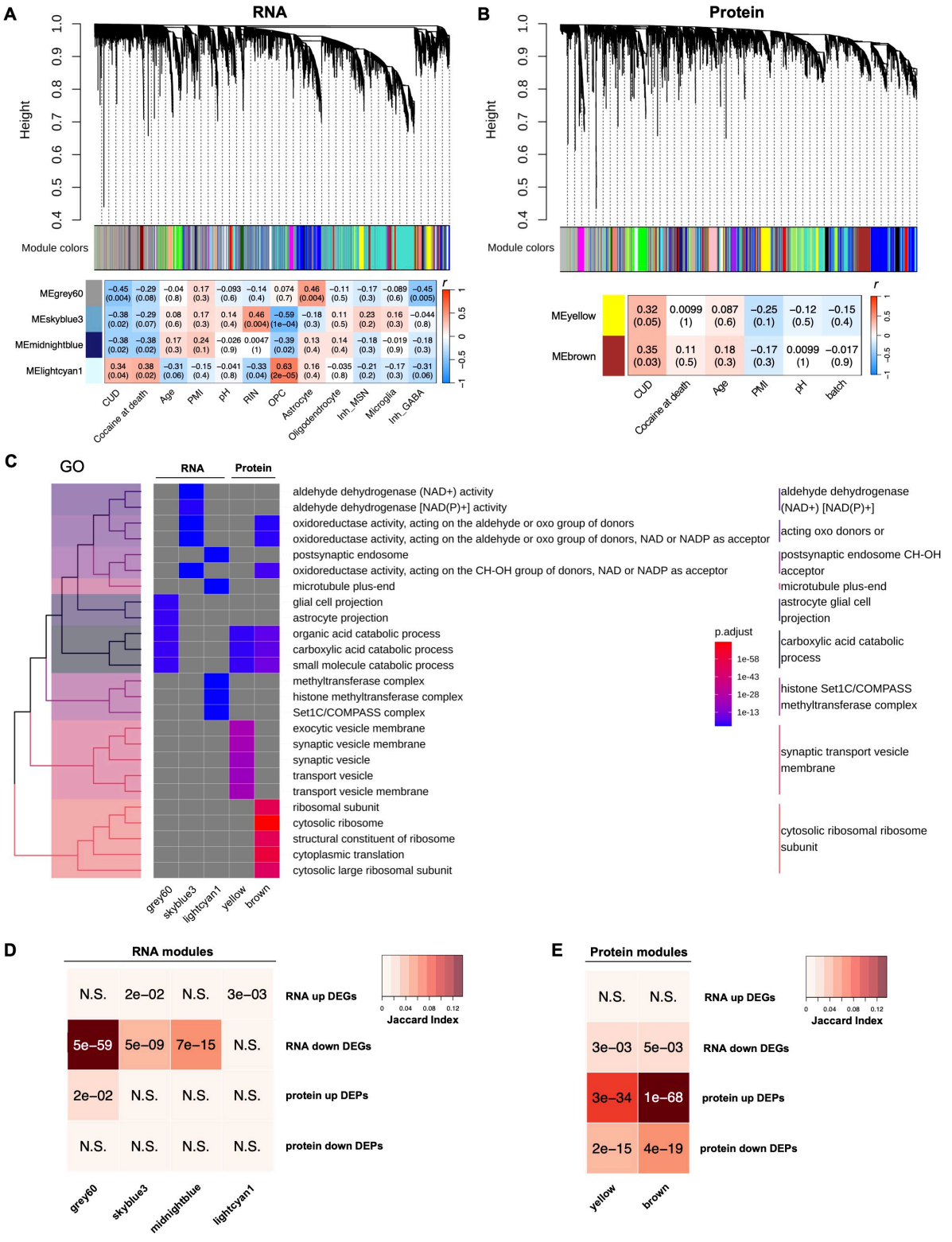


Figure 4.2 - Identification and functional characterization of CUD-associated RNA and protein co-expression networks

(A and B) Dendrogram resulting from weighted co-expression network analysis (WGCNA) in (A) RNA-seq and (B) proteomics datasets. Module eigengene (ME) correlation with CUD, available covariates, and cell type estimates from deconvolution analysis are shown for CUD-associated co-expression modules. Panels contain color-coded Pearson correlation coefficient (*r*). Significance of correlation is shown in brackets. PMI, postmortem interval; pH, postmortem brain tissue pH value; RIN, RNA integrity number; OPC, oligodendrocyte progenitor cell; Inh_MSN, cell type estimates for DRD1- and DRD2-expressing medium spiny neurons (MSNs); Inh_GABA, cell type estimates

for other GABAergic interneurons of the VS not included in the DRD1- and DRD2-expressing MSN clusters; batch, proteomics processing batch. (C) Functional characterization of co-expressed genes in RNA and protein co-expression modules in CUD. Statistically significant results from GO enrichment analysis results of co-expression module genes were visualized using a treeplot. (D and E) Analysis of overlap between up- and downregulated DEGs and DEPs and co-expression modules for (D) RNA co-expression and (E) protein co-expression module genes. Panels contain *p* values from a Fisher test indicating the significance of overlap. DEG, differentially expressed gene from RNA-seq (q < 0.25); DEP, differentially expressed protein (q < 0.25); N.S., not significant.

The association with factor 10 was even stronger for cocaine at death ($r = 0.47$, $p = 0.002$) suggesting that factor 10 might particularly reflect the intoxication state of CUD. Factor 10 was further associated with RNA-based cell type estimates for astrocytes ($r = -0.38$, $p = 0.02$), GABAergic inhibitory neurons ($r = -0.39$, $p = 0.01$), and oligodendrocytes ($r = 0.32$, $p = 0.05$). To further analyze the biological processes represented by CUD-associated factor 10, we inspected the weights of individual miRNAs, RNAs, and proteins on this factor (Table S4.4D). hsa-miR-30c-2-3p and hsa-miR-101-3p were identified as the miRNAs with the strongest negative and positive weights, respectively (Figure 4.3D). At the RNA level, the top features were electron transport chain-associated gene *ETFRF1* for negative weights and nicotinic acetylcholine receptor modulator *LY6H* for positive weights (Figure 4.3E). At the protein level, UBA2 related to protein SUMOylation and the GTPase RAB14 involved in vesicle trafficking displayed the largest absolute negative and positive weights, respectively (Figure 4.3F). Using weight information from RNA and protein datasets, pre-ranked gene set enrichment analysis (GSEA) was performed to identify biological pathways associated with factor 10 (Table S4.4E). Largest functional pathway modules after GO term clustering were related to neuronal morphology and oxidative phosphorylation (Figure 4.3G). Among significant results with positive GSEA normalized enrichment score (NES) at the protein level, synaptic signaling pathways emerged, confirming results from the WGCNA analysis of the proteomic dataset. A second highly connected pathway cluster consisting of metabolic GO terms related to oxidative phosphorylation was specific to protein results with negative NES. Thus, while previous pathway enrichment analyses of DEGs, DEPs, and WGCNA module genes suggest alterations in fatty acid metabolism and oxidoreductase enzymes, MOFA identified additional alterations in oxidative phosphorylation depicting a shared downstream process of the previously identified metabolic processes.

4.7.4 snRNA-seq identifies the cell type specificity of transcriptional changes and suggests astrocytes and medium spiny neurons as important cell types in CUD

Our bulk-level analyses in the VS provide evidence for metabolic, ribosomal, and synaptic changes in CUD. However, it remains unclear to what extent these findings are specific to or driven by individual cell types. To identify potential cell type-specific transcriptomic changes in CUD, we performed snRNA-seq in a subset

of n = 16 individuals (eight CUD cases, eight Ctrl individuals; Table S4.1). In our dataset of n = 20,759 single nuclei, we identified 12 distinct cell type clusters in the VS (Figures 4.4A, 4.4B, and S4.3A–S4.3G). Major neuronal cell types of the VS include DRD1- and DRD2-expressing medium spiny neurons (MSNs; D1-MSNs/D2-MSNs) as well as several non-MSN GABAergic interneuron populations. Investigation of marker gene expression in neuronal clusters confirmed the presence of known interneuron populations of the striatum (Garma et al., 2024; Gokce et al., 2016), including PTHLH-, VIP-, SST-, NPY-, and CCK-expressing interneurons (GABAergic-1); a larger TAC3-expressing population (GABAergic-2); as well as small MSN subpopulations such as ADARB2-positive D1-MSNs (GABAergic-1) and DRD2-expressing HTR7-MSNs (GABAergic-3; Figure S4.3H).

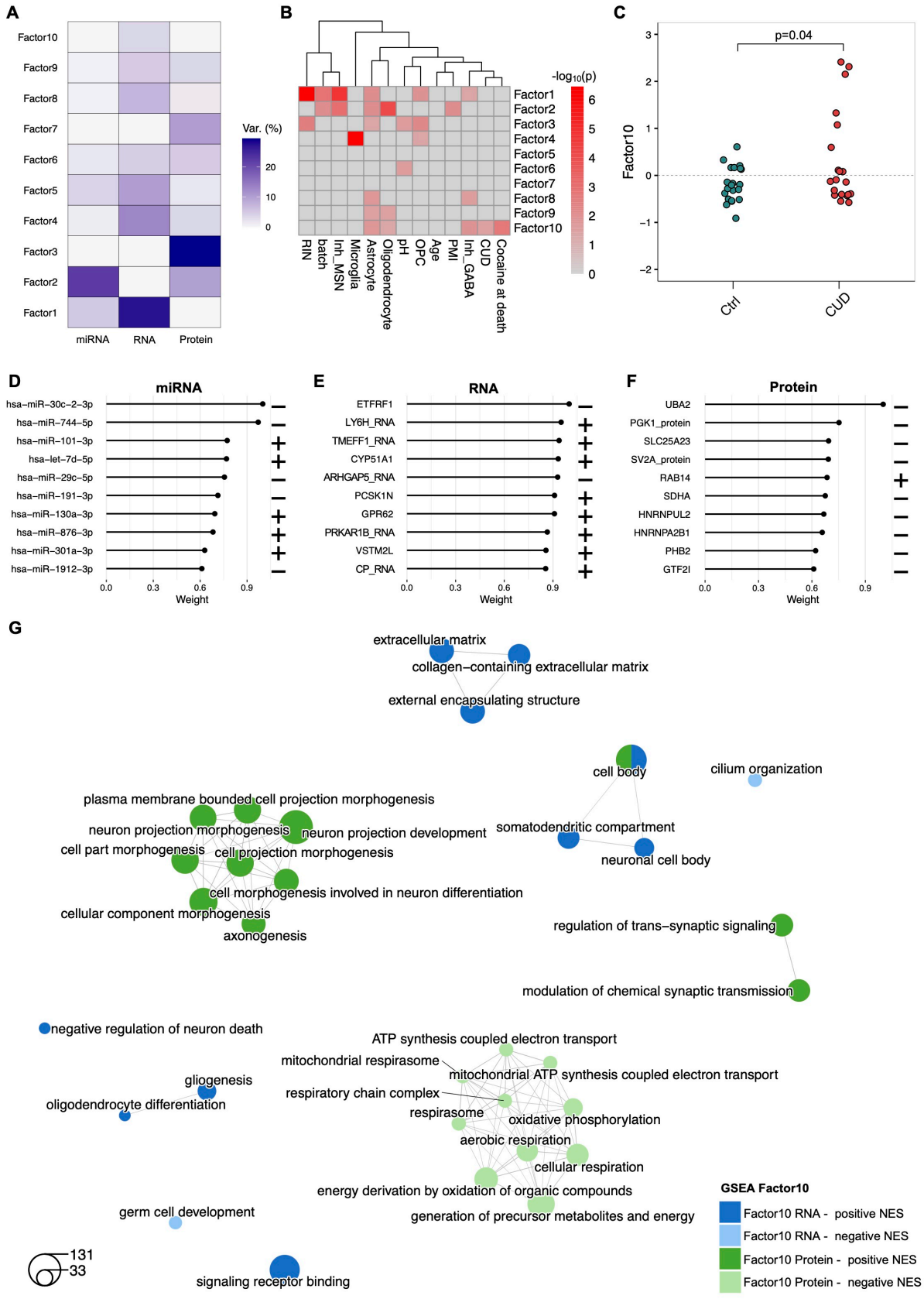


Figure 4.3 - Integrative bulk-level analysis of miRNA-seq, RNA-seq, and proteomics data by multi-omics factor analysis reveals a CUD-associated latent factor enriched for synaptic and metabolic genes

(A) Variance explanation per MOFA factor for each omics dataset, miRNA-seq (all 1,542 miRNAs), RNA-seq (top 4,270 variable RNAs), and proteomics (all 4,270 proteins). (B) Correlation of factors

with CUD, covariates, and cell type estimates from cell type deconvolution analysis in RNA-seq data. Significance of correlation as $-\log_{10}(p \text{ value})$ is shown in the heatmap. (C) Comparison of sample loadings on factor 10. *p* value from a Wilcoxon test on factor 10 loadings is shown. (D–F) Top 10 features, i.e., (D) miRNAs, (E) RNAs, and (F) proteins with strongest absolute weights on CUD-associated factor 10. "+" indicates positive and "–" indicates negative weight. (G) Gene set enrichment analysis (GSEA) results for positive and negative RNA (blue) and protein weights (green) on factor 10 identifies functional pathway modules among statistically significant GO terms (FDR $q < 0.05$).

Further, glial cell types such as astrocytes, oligodendrocytes, oligodendrocyte progenitor cells (OPCs), and microglia were identified in our dataset (Figures 4.4A and 4.4B). To determine DEGs associated with CUD, we performed CUD vs. Ctrl differential expression analysis in each major cell type cluster identifying a total of 653 DEGs ($|\log2FC|>0.5$, $q < 0.001$; Table S4.5A; Figure 4.4C). These include well-described immediate-early genes such as *JUN* and *FOSB* that have been repeatedly described to be induced in human brain and rodent brain following cocaine exposure (Bannon et al., 2014; Hope et al., 1992; Salery et al., 2021). The strongest CUD-associated expression deregulation was found in D1-MSNs (296 DEGs), D2-MSNs (200 DEGs), and astrocytes (128 DEGs; Figure S4.3I), which is well reflected by cell type-specific DEG patterns in a rodent model of repeated cocaine intake (Phillips et al., 2023). While the large number of DEGs identified for these cell types could be a consequence of the cluster size and the associated stronger statistical power in our dataset, significantly fewer DEGs were found in similarly sized or larger clusters such as oligodendrocytes (48 DEGs) or microglia (91 DEGs), suggesting D1-MSNs, D2-MSNs, and astrocytes as particularly deregulated cell types in CUD. As RNA expression data were available at the bulk- and single-cell level in our study, we compared CUD-associated DEG patterns between bulk RNA-seq and snRNA-seq (differential expression [DE] analysis across all clusters) using RRHO. Here, a strong convergence of results was observed (Figure S4.3J) confirming robust overlap between DEG signatures in bulk-level RNA-seq and cluster-ignorant snRNA-seq analysis. To investigate cell type-specific pathway deregulation, a GO enrichment analysis was performed using DEGs from each of the seven major cell type clusters (Figure 4.4D; Table S4.5B). Significant GO term clusters conserved across neuronal and glial cell types include "structural constituent of ribosome" (D1-MSN, D2-MSN, astrocyte), oxidative phosphorylation pathways (D1-MSN, astrocyte, oligodendrocyte), and "ubiquitin-like protein ligase binding" (D1-MSN, oligodendrocyte). Ribosomal and oxidative phosphorylation changes might be a conserved feature of CUD across striatal brain regions as they also emerged in a recent cell type-specific analysis in the CN (Zillich et al., 2025). Further, we found significant pathway modules that were specific to either neurons or glial clusters, such as immunological pathways related to T cell receptor and major histocompatibility complex (MHC) binding (microglia, OPC), and mRNA 5′-UTR binding (D1-MSN, D2-MSN). As ribosomal and oxidative phosphorylation pathways were reported in existing literature (Zhou et al., 2011; Zillich et al.,

2025), have emerged as main findings in our bulk-level analyses, and were conserved across neuronal and non-neuronal clusters from snRNA-seq, we were interested in the direction of DEG patterns for these biological processes in individual cell types. We evaluated log2FC and *p* values of genes related to ribosomal and electron transport chain (ETC) pathways by inspecting genes involved in "structural constituent of ribosome" and "electron transfer activity" GO terms. For both ribosomal and ETC genes, significant upregulation was detected in neurons, especially in D1-MSNs and D2-MSNs, while the same genes were significantly downregulated in glial cell types, most prominently in astrocytes and oligodendrocytes (Figures 4.4E and 4.4F). Cell type-specific analysis of gene expression in CUD thus suggests inverse deregulation patterns between neurons and glia cells within genes related to the ribosome and oxidative phosphorylation.

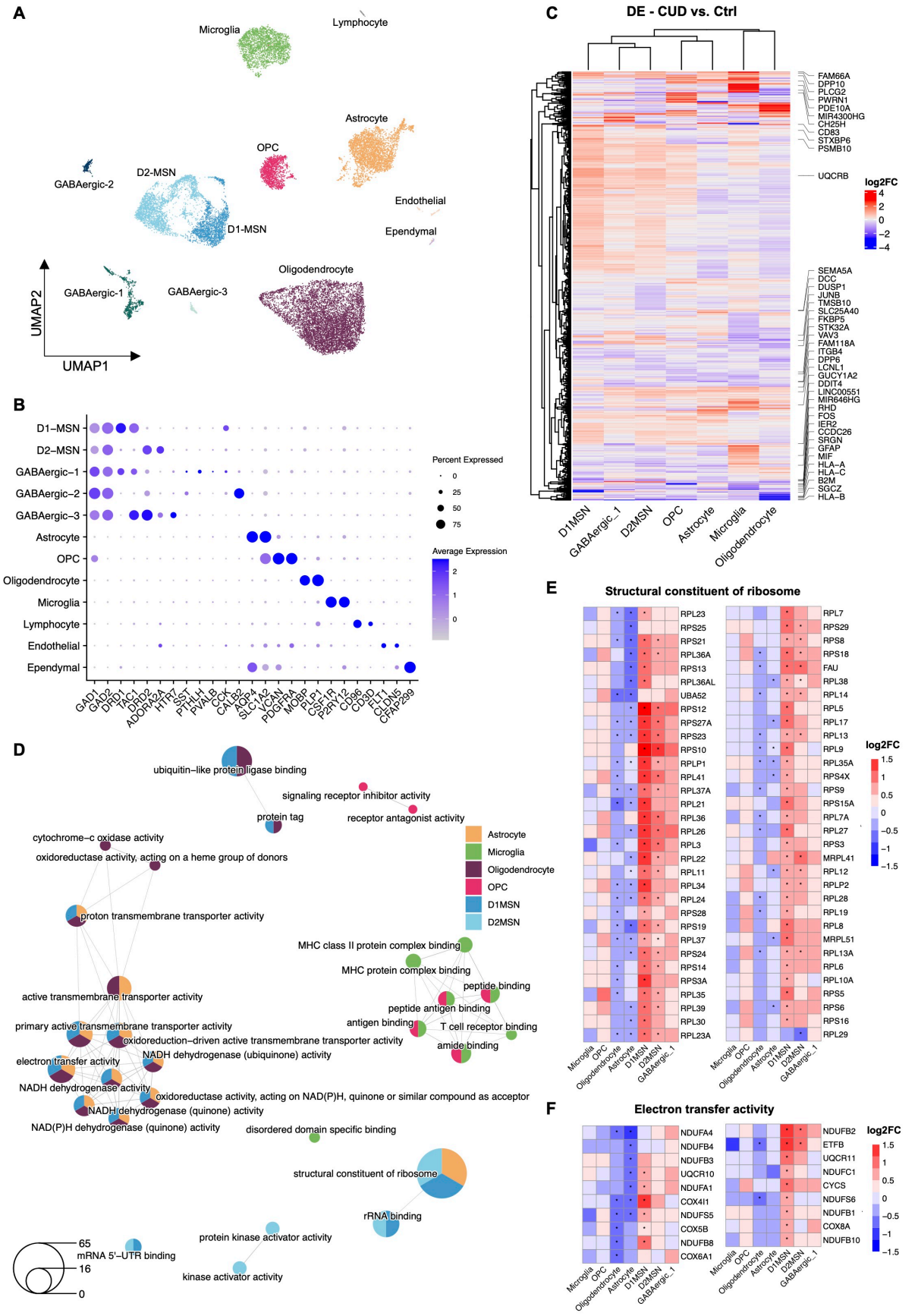


Figure 4.4 - snRNA-seq of the VS identifies cell type-specific transcriptomic deregulation patterns in CUD

(A) Uniform manifold approximation and projection (UMAP) representation of the transcriptomic profiles of n = 20,759 single nuclei from n = 8 individuals with and n = 8 individuals without CUD

identifies 12 cell type clusters in the VS. (B) Annotation of cell types based on expression levels of known cell type marker genes. (C) Heatmap of top differentially expressed genes (DEGs) in the seven major cell types of the VS showing the $n = 653$ DEGs characterized by $|\log2FC|>0.5$ and $q < 0.001$. Top deregulated genes per cell type based on expression log2FC are highlighted by gene name. (D) Statistically significant results from GO enrichment analysis (FDR $q < 0.05$) of DEGs ($|\log2FC|>0.5$, $q < 0.001$) from major cell types. Heatmap of RNA log2FCs of pathway-defining genes for pathways characterized by significant enrichment of DEGs in both neuronal and glial cell types. (E and F) Genes corresponding to GO terms (E) structural constituent of ribosome and (F) electron transfer activity were investigated. *DEG with $q < 0.001$.

4.7.5 Cell type-specific co-expression analysis identifies gene-network alterations in astrocytes and striatal MSNs that are conserved across human CUD and a rodent model of repeated cocaine intake

Consistent with the analysis approach in bulk-level datasets, we aimed to identify and functionally characterize transcriptional co-expression networks in the single-nuclei dataset. We first performed a cell type-specific co-expression network analysis in our human snRNA-seq data using high-dimensional WGCNA (hdWGCNA) (Morabito et al., 2023). Module prioritization (Table S4.6A; STAR Methods) resulted in eight cell type-specific co-expression modules characterized by significant downregulation in CUD (differential module eigengene $\log2FC < 0$, $q < 0.05$; Figures S4.4A–S4.4D; Table S4.6A). Cell type-specific co-expression modules were identified in astrocytes, MSNs (Inh_MSN), and GABAergic interneurons clusters (Inh_GABA; Table S4.6B). Co-expression module genes were significantly enriched in cell type-specific DEGs ($|\log2FC|>0.5$, $q < 0.001$), especially in downregulated DEGs, further supporting the downregulation of the identified cell type-specific co-expression networks in CUD (Figure S4.4E). Pathway enrichment analysis in astrocytic co-expression networks highlights deregulation of glutamatergic synapses and ion transport processes as well as aberrant glutamate and fatty acid metabolism in CUD (Astrocyte-M12, Astrocyte-M14; Figures S4.5A and S4.5B; Table S4.6C). In line with this, we found important regulators of glutamatergic signaling such as glutamate dehydrogenase *GLUD1* (Astrocyte-M12; Figure S4.4F), glutamate transporters *SLC1A2* and *SLC1A3* (Astrocyte-M14; Figure S4.4F), as well as metabotropic glutamate receptor *GRM3* (Astrocyte-M14; Figure S4.4F) among module hub genes in the astrocyte-specific CUD-associated co-expression modules. In neuron-specific co-expression modules, we identified metabolic pathways related to nucleoside and ketone body metabolism, ion transport processes, and GABAergic signaling (Figures S4.5C and S4.5D; Table S4.6C). Module hub genes include several ionotropic neurotransmitter receptor genes such as GABA-A receptor subunits *GABRB1* and *GABRB3*, potassium channel gene *KCNQ5*, calcium channel subunit *CACNA1E* (all Inh_MSN-M7; Figure S4.4F), sodium-calcium exchanger *SLC8A1*, and *SLC22A17* involved in iron transport (both Inh-MSN-M2; Figure S4.4F).

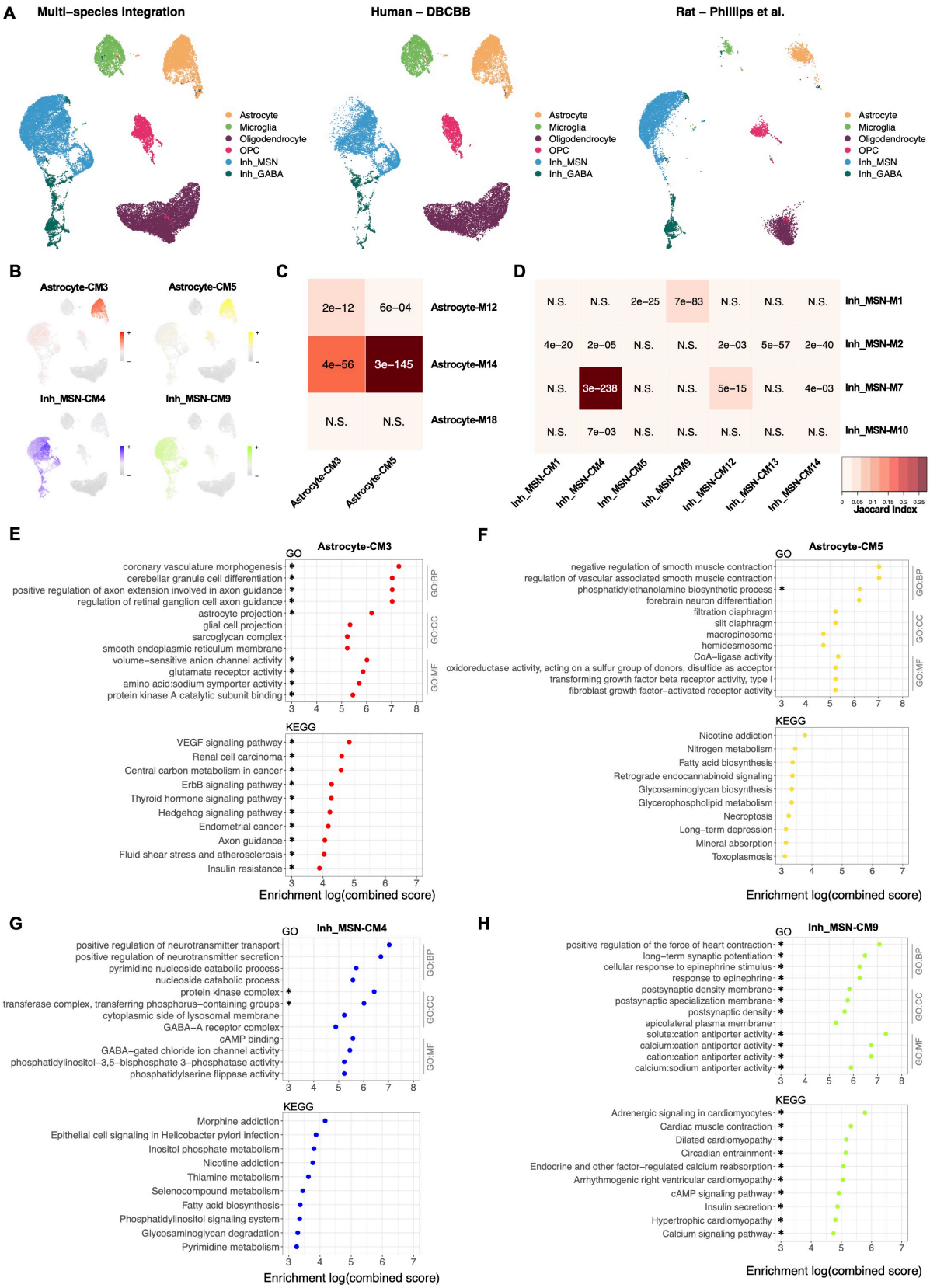


Figure 4.5 - Integrative analysis of snRNA-seq datasets from human CUD and a rat model of repeated cocaine intake reveals consensus co-expression modules across species
(A) UMAP representations of the integrated snRNA-seq dataset containing n = 20,492 human and n = 11,288 rat nuclei of ventral striatum cell major types. DBCBB, Douglas Bell Canada Brain Bank, rat dataset from Phillips et al. (2023) (B) Expression of module eigengenes from consensus co-expression modules with significant DME in CUD and cell type-specific expression. (C and D)

Analysis of the relationship between human-only and consensus co-expression modules in (C) astrocyte and (D) Inh_MSN clusters measuring the overlap of co-expression module genes. Jaccard index and *p* values from Fisher test are shown. N.S., not significant. (E–H) GO and KEGG enrichment analysis results for consensus co-expression modules (E) Astrocyte-CM3, (F) Astrocyte-CM5, (G) Inh_MSN-CM4, and (H) Inh_MSN-CM9 characterized by strongest module gene overlap with co-expression modules in the human dataset. For GO terms, the top four pathways from each ontology (GO Biological Process [GO:BP], GO Cellular Component [GO:CC], and GO Molecular Function [GO:MF]) are shown. *FDR q < 0.05.

To investigate potential conservation patterns of network changes across human CUD and a repeated cocaine intake paradigm in rats, we performed consensus hdWGCNA in the n = 20,759 human nuclei and n = 11,288 nuclei of the VS (nucleus accumbens) from male Sprague-Dawley rats undergoing 7 days of cocaine exposure (Mews et al., 2023) (Figure 4.5A). For MSNs and astrocytes, the cell types that have been most prominently implicated in human CUD based on DE and hdWGCNA analyses and also showed the largest number of DEGs in the rodent dataset, we found seven (Inh_MSN-CM1, Inh_MSN-CM4, Inh_MSN-CM5, Inh_MSN-CM9, Inh_MSN-CM12, Inh_MSN-CM13, Inh_MSN-CM14) and two (Astrocyte-CM3, Astrocyte-CM5) cell type-specific CUD-associated consensus co-expression modules, respectively (Tables S4.6D and S4.6E). In astrocytes, a strong overlap was observed between human module genes (Astrocyte-M14) and consensus module genes (Astrocyte-CM3, Astrocyte-CM5; Figures 4.5B and 4.5C). Consistent with the hdWGCNA analysis in human CUD, the two astrocyte-specific consensus modules displayed negative differential module eigengene (DME) results indicating downregulation of the consensus networks in CUD (Table S4.6D). Astrocyte-specific consensus modules were enriched for pathways involved in axon guidance, fatty acid metabolism, and glutamatergic signaling (Figures 4.5E and 4.5F; Table S4.6F) supporting pathway results from human modules. In neurons, the strongest overlap was observed between human module Inh_MSN-M7 and consensus module Inh_MSN-CM4, which was also downregulated in CUD (DME log2FC = −0.59; Figures 4.5B and 4.5D; Table S4.15). At the pathway level, GABAergic signaling as well as nucleoside and fatty acid metabolism changes were among the significant findings in neuron-specific consensus module Inh_MSN-CM4 (Figure 4.5G; Table S4.6D) in line with results from the human module Inh_MSN-M7. In consensus module Inh_MSN-CM9, significant pathways were related to ion transport processes involved in calcium signaling (Figure 4.5H; Table S4.6F). Notably, KEGG enrichment analysis of consensus co-expression module genes revealed multiple SUD-related pathways among the most significant associations such as morphine addiction and nicotine addiction in both astrocytes and MSNs, confirming the presence of addiction-relevant genes in consensus co-expression modules (Figures 4.5E–4.5H). Results from the consensus network analysis thus indicate a set of deregulated biological processes including fatty acid metabolism and glutamatergic signaling conserved between human CUD and a rat model of repeated cocaine intake.

4.7.6 Altered astrocyte-neuron crosstalk in human CUD affects glutamatergic signaling and communication via cell adhesion molecules

Most prominent changes in CUD-associated gene expression and co-expression networks were found in astrocytes and MSNs of the VS. To better understand the consequences of transcriptional deregulation in these cell types, we investigated cell-cell crosstalk by performing CellChat in astrocyte, D1-MSN, and D2-MSN clusters. We investigated differential activity of ligand-receptor (LR) interaction pairs in CUD and found upregulation of 16 and downregulation of 106 LR pairs in CUD (Table S4.7).

Upregulated signaling in CUD was affecting secreted signaling factors such as glycoprotein *SLIT2* or Eph receptor ligand *EFNA5*, originating from D1-MSNs (Figure 4.6A). D1-MSNs and D2-MSNs were also the main receiving cell type of upregulated signaling in CUD, indicating limited neuron-astrocyte crosstalk through upregulated LR pairs. In contrast, neuron-astrocyte crosstalk was particularly present among downregulated LR pairs, with the strongest expression changes of ligands detected in astrocytes and D2-MSNs (Figure 4.6B). The astrocytic ligand expression profile was characterized by downregulation of glutamatergic signaling genes (*SLC1A2*, *SLC1A3*) in line with results from astrocyte-specific co-expression modules and supported by consistent findings of *SLC1A2* downregulation in rodent models of cocaine addiction (Niedzielska-Andres et al., 2021). Further, among downregulated ligands on astrocytes, we found several cell adhesion molecules, including *NRXN1* and *CADM1*, that have an important role in securing the structural integrity of tripartite synapses (Hillen et al., 2018). Both D1-MSNs and D2-MSNs were receivers of altered glutamate signaling from astrocytes as subunits of AMPA (*GRIA3*, *GRIA4*) and kainate receptors (*GRIK2*, *GRIK3*) as well as metabotropic glutamate receptor *GRM7* were downregulated in MSNs. Further, D2-MNSs were the main receivers of differential cell-cell adhesion signaling as they expressed reduced levels of NRXN interaction partners *NLGN1* and LRRTM family members *LRRTM3* and *LRRTM4*. Thus, our analysis of differential LR pair expression in CUD suggests altered astrocyte-neuron crosstalk related to the deregulation of glutamatergic and cell-cell adhesion signaling. Interestingly, the same set of glutamatergic and cell-cell adhesion signaling genes have been identified as module hub genes of hdWGCNA co-expression modules Astrocyte-M14 and Inh_MSN-M7 (Figure S4.4F), which were both downregulated in CUD and displayed conservation patterns across species. CellChat analysis thus confirms findings from co-expression analysis and further outlines altered crosstalk between astrocytes and MSNs as an important hallmark of the VS in CUD.

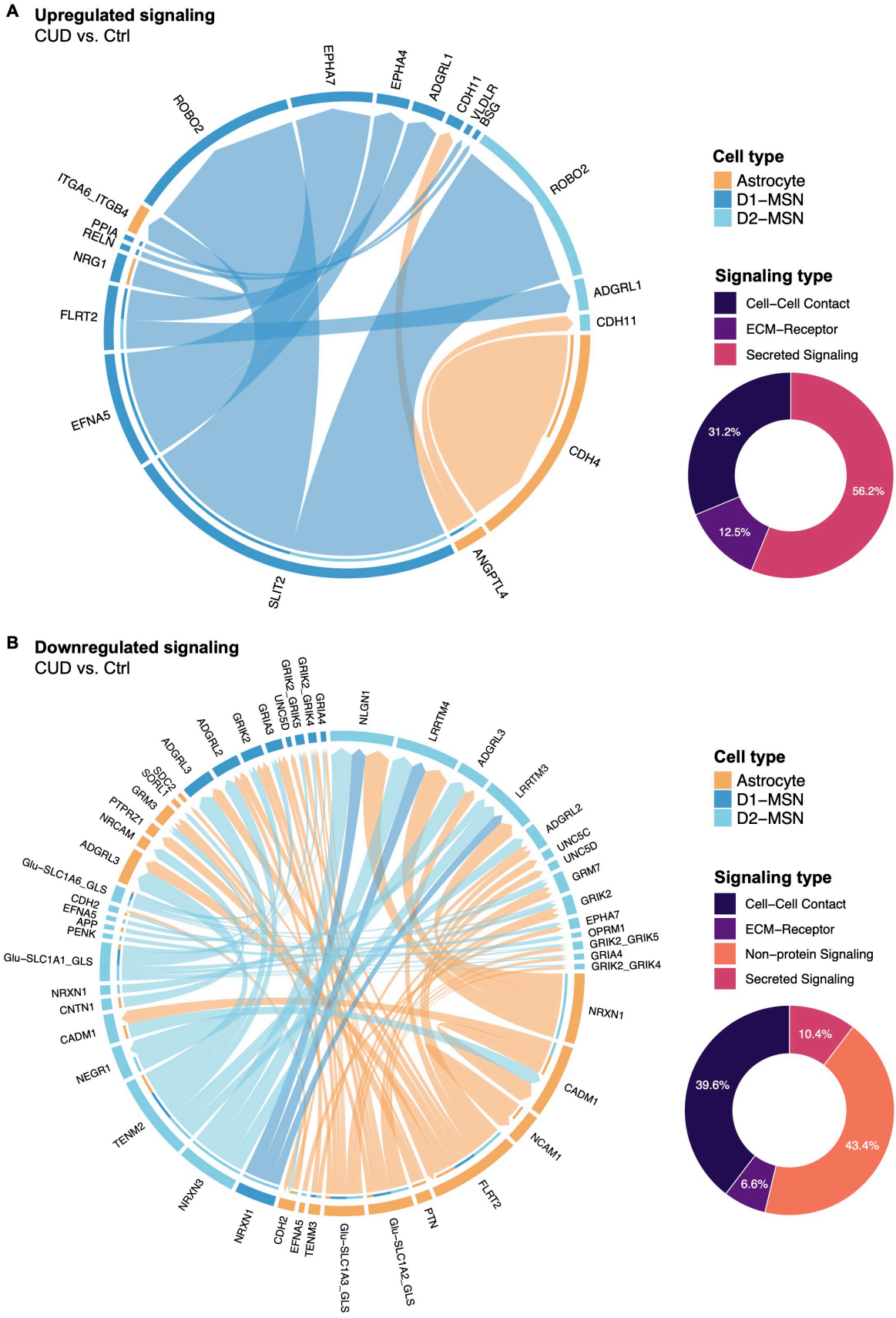


Figure 4.6 - Astrocyte-neuron crosstalk in CUD is characterized by aberrant glutamatergic and cell-cell adhesion signaling
(A and B) Circos plot for significantly (A) upregulated and (B) downregulated LR pairs in astrocytes (orange), D1-MSNs (blue), and D2-MSNs (light blue). Arrows indicate the directionality of signaling

from a ligand expressed in the sender cell (starting point of the arrow) to its receptor expressed in the receiver cell type (arrowhead). Percentage of signaling types among upregulated LR pairs are shown in the donut plot. Cutoffs for significant DE of LR pairs were a minimum of 5% change in ligand expression ($|log2FC > 0.07|$) and receptor DE into the same direction, both at 5% FDR significance.

4.8 Discussion

The present study depicts an integrative multi-omics analysis framework for characterizing neurobiological changes in the human VS in CUD. To our knowledge, our study is the first in a psychiatric phenotype that performed multi-omics integration of microRNA-seq, RNA-seq, and proteomics datasets together with cell type-specific analyses by snRNA-seq in postmortem human brain tissue. Our study design enables the identification of molecular changes in CUD at different levels of biological regulation—starting from miRNAs that act on the RNA level up to the protein level. This allowed us to provide novel insights into across-omics conservation patterns of molecular deregulation in CUD, for instance, involving fatty acid metabolism. By performing additional analyses at single-cell resolution, we showed that bulk-level results, such as ribosomal and oxidative phosphorylation changes, were also reflected in the single-nuclei dataset. At the same time, we gained important insights into directional effects of transcript deregulation patterns in glial compared to neuronal cell types, highlighting the value of cell type-specific analyses. Finally, we show the advantage of integrating rodent model and human datasets in substance use disorders, where we found converging evidence for altered glutamatergic signaling in CUD that was also confirmed by an astrocyte-neuron crosstalk analysis.

Pathway associations with CUD from bulk-level analyses include immunological, ribosomal, synaptic, and metabolic changes related to fatty acid metabolism and oxidative phosphorylation. This confirms findings from previous analyses of miRNA, RNA transcriptomic, and proteomic studies in CUD (Mews et al., 2023; Poisel et al., 2023; Ribeiro et al., 2017; Tondo et al., 2021; Vaillancourt, Chen, et al., 2021; Zhou et al., 2011; Zillich et al., 2024; Zillich et al., 2025). Importantly, our study provides the first insights into the inter-relationship between layers of biological regulation in CUD by identifying across-omics conservation (for instance, related to fatty acid metabolism as indicated by WGCNA co-expression analysis). This underpins the deregulation of fatty acid metabolism as a metabolic feature of the CUD brain and suggests that these metabolic alterations previously observed at the RNA level (Pati et al., 2019; Zillich et al., 2024) also extend to the protein level. While convergent evidence across omics analyses was observed, we also found several biological processes exclusively detected in single-omics analyses. From this observation, we conclude that integrative multi-omics analyses depict an important approach for a better understanding of molecular

changes in complex phenotypes such as CUD, as single-omics analysis might miss important disease associations.

A frequent criticism of gene expression studies is that transcript levels are poorly predictive for protein levels, thus limiting the interpretability of findings. Our transcriptome-proteome correlation analysis revealed an overall moderate correlation (r = 0.43) between RNA and protein levels in the VS, which is well in line with results from a study investigating schizophrenia-associated molecular changes in the prefrontal cortex that also observed a correlation of r = 0.43 between transcriptome and proteome (Luo et al., 2024). Importantly, we found the strongest positive correlation between RNA and protein levels for genes involved in neuronal function and synaptic signaling such as *PDE10A*, *SCN4B*, and *FKBP5*, suggesting that the assumption of RNA levels serving as an indicator for protein levels holds true for these frequently studied genes in psychiatric disorders. Strong negative RNA-protein correlation was detected for genes involved in oxidative phosphorylation. While this might be a true biological effect, it could also be a consequence of insufficient mitochondrial protein solubilization during brain tissue lysis.

In cell type-specific DEG analyses, we found inverse deregulation patterns between neuronal and glial cell types related to ribosomal and oxidative phosphorylation pathways. Alterations of these pathways in CUD is in line with results from previous postmortem human brain studies (Zhou et al., 2011; Zillich et al., 2025). In a cell type-specific multi-omics study of CUD using snRNA-seq and snATAC-seq, we found similar ribosomal gene deregulation patterns in the CN (Zillich et al., 2025). One explanation for a neuron-specific upregulation of ribosomal genes in CUD could be an increased demand for local translation, particularly in neurons, as ribosomes are abundant in dendrites and synapses where they are required for the local synthesis of proteins involved in neurotransmission and synapse structure (Sun et al., 2021). Our finding of increased ribosomal protein levels based on bulk-level proteomics in the VS further supports the hypothesis of an increased ribosomal demand in CUD. Further, DE of oxidative phosphorylation genes was detected in a bulk-level RNA-seq study of the hippocampus in CUD (Zhou et al., 2011). While the authors report on the overall downregulation of oxidative phosphorylation genes, we show that this finding is cell type dependent in the VS, with downregulation patterns observed for glial cell types, while the same genes were upregulated in neurons. We further found fatty acid metabolism changes to be especially prominent among astrocyte-specific co-expression modules, highlighting the value of cell type-specific analyses to better understand findings that have been previously reported in bulk-level analyses. Follow-up studies are required to investigate whether ribosomal and oxidative phosphorylation changes depict direct effects of cocaine exposure or are compensatory effects induced by long-term cocaine intake.

Interestingly, we did not observe a considerable fraction of neurotransmission or synaptic plasticity pathways in the differential transcript expression analysis of bulk- or snRNA-seq datasets of the VS. In contrast, bulk-level RNA-seq analysis of the prefrontal cortex subregion Brodmann area 9 (BA9) in a subset of $n = 25$ individuals from our Douglas Bell Canada Brain Bank (DBCBB) cohort revealed strong transcriptional changes for synaptic signaling genes in CUD (Zillich et al., 2024). Also, in the snRNA-seq/snATAC-seq study of CUD in the CN, a brain region involved in compulsive drug intake patterns harboring similar cell types as the VS, we found consistent alterations in synaptic and ion channel signaling across transcriptomic and chromatin accessibility datasets (Zillich et al., 2025). At the same time, metabolic findings related to fatty acid metabolism or oxidative phosphorylation were consistently associated with CUD across the investigated brain regions. We hypothesize that this observation could reflect an aspect of late-stage CUD where the VS has undergone profound neuroadaptations in response to chronic cocaine intake but resides in an anaplastic state characterized by suppression of long-term depression and an inability of synaptic reshaping (Kasanetz et al., 2010).

A strong advantage of rodent models in addiction research is controlled experimental conditions, allowing mechanistic studies on drug effects by systematic exclusion of confounding factors. To complement the study of postmortem brains, which is associated with strong inter-individual heterogeneity, with a controlled laboratory experiment, we performed integration of our human snRNA-seq data with snRNA-seq from a rat model of repeated cocaine intake. Here, we deciphered conserved CUD-associated gene networks in the VS. A prominent network finding was glutamatergic signaling alterations that were conserved across human and rat datasets as they emerged in human hdWGCNA, consensus hdWGCNA, and astrocyte-neuron crosstalk analysis. Craving and relapse, two key symptoms of CUD, have been shown to correlate with aberrant glutamate signaling in the VS (Cornish & Kalivas, 2000) and our study identified astrocytes and MSNs as important cell types that may contribute to this observation. While D1- and D2-MSNs have been recognized early on as important cell types for addiction research, there is increasing evidence for a key role of astrocytes in CUD (Kruyer & Scofield, 2021; Wang et al., 2022). For instance, the neurovascular effects of cocaine characterized by cerebral blood flow reduction were shown to be dependent on astrocytic calcium signaling, and chemogenetic inhibition of astrocytes prevented cerebral vasoconstriction following a cocaine challenge in rodents (Du et al., 2024). Regarding glutamatergic signaling, glutamate transporter GLT-1 (SLC1A2) was consistently downregulated in NAc astrocytes of rats repeatedly exposed to cocaine, and experimental normalization of GLT-1 levels reduced reinstatement of cocaine seeking (Knackstedt et al., 2010; Niedzielska-Andres et al., 2021; Reissner et al., 2014). Our study confirms a downregulation of *GLT-1* in in human CUD astrocytes while also suggesting

ionotropic and metabotropic glutamate receptor subunits to be downregulated in D1- and D2-MSNs. Another important finding of the altered astrocyte-neuron crosstalk in CUD were changes in cell-cell adhesion dynamics that could further enhance abnormal glutamatergic signaling, as reduced NRXN-NLGN interaction disrupts the structural integrity of tripartite synapses, thereby additionally impeding glutamate homeostasis (Trotter et al., 2021; Walker et al., 2020). Furthermore, the glutamatergic imbalance in the VS depicts a possible link to the oxidative phosphorylation changes for which we observed different directional effects in glia and neurons. For instance, conversion of glutamate to α-ketoglutarate was shown to fuel the tricarboxylic acid cycle in brain mitochondria, thereby temporarily increasing ATP production via oxidative phosphorylation changes (Panov et al., 2009). Further studies are required to disentangle the direct and adaptive effects of aberrant glutamatergic signaling with the aim to investigate the modulation of glutamate system genes as a potential pharmacotherapy in CUD.

4.8.1 Limitations of the study

While aiming for the largest sample size possible, a limitation of the present study is the relatively small discovery cohort, which might not represent the full spectrum of the CUD phenotype. Our cohort was selected to be homogeneous in sex and ancestry, which is advantageous for statistical analyses in postmortem cohorts of limited size. However, this results in a lower generalizability of the results to the general population, which underlines the need for analyses in more diverse cohorts. Strong inter-individual heterogeneity depicts a general phenomenon in CUD, as cocaine is frequently consumed together with other drugs of abuse such as alcohol, cannabis, or opioids (Stiltner et al., 2023). It thus cannot be excluded that additional factors such as exposure to other drugs of abuse, medication prior to death, and cause or manner of death might interfere with CUD-associated DE signatures. With the available sample size, we were not able to adjust for all potential influences on miRNA, RNA, and protein expression levels, but, in larger cohorts, the inclusion of additional covariates together with subgroup analyses could be useful to address this problem. For the interpretation of findings, it needs to be considered that the cross-sectional design does not allow us to distinguish between cumulative effects of cocaine exposure and compensatory neuroadaptations. The sample size and the number of retained nuclei in the snRNA-seq dataset depicts another limitation. Larger sample sizes would allow the use of advanced statistical methods for DE modeling, such as negative binomial mixed models (He et al., 2021), which are less prone to false positives than the Wilcoxon rank-sum test, while preserving the hierarchical but granular architecture of single-nuclei datasets. The analysis of the snRNA-seq dataset also revealed a strong underrepresentation, particularly of striatal non-MSN GABAergic neuron populations such as cholinergic interneurons. This may be a

consequence of a particular susceptibility of certain cell types to the tissue homogenization and nuclei isolation procedure, resulting in a dataset that might not be fully representative of the cellular complexity of the human VS.

Future studies should focus on the identification and validation of disease mechanisms based on multi-omics analyses, for instance by identifying master regulators of CUD-associated transcriptional changes in individual cell types (Zillich et al., 2025). Further, multi-omics integration and conservation analyses between human CUD and rodent models capturing addiction-like criteria should be performed to characterize conserved patterns of molecular deregulation, thereby addressing the inherent limitations of the two study types (Spanagel et al., 2024). In summary, such multi-omics-to-mechanism studies provide a powerful analytical framework for the identification of disease mechanisms and potential new therapeutic targets in CUD.

4.9 Resource availability

4.9.1 Lead contact

Requests for further information, resources, and/or reagents should be directed to the lead contact, Stephanie Witt (stephanie.witt@zi-mannheim.de).

4.9.2 Materials availability

The present study did not generate new materials.

4.9.3 Data and code availability

- Raw sequencing (miRNA-seq, RNA-seq, snRNA-seq) and proteomic datasets are deposited at the European Genome-phenome Archive (EGA), which is hosted by the European Bioinformatics Institute (EBI) and the Centre for Genomic Regulation (CRG), under study accession numbers EGAS50000000623 and EGAS00001007945.

- All original code used for data analysis and figure preparation is available in a GitHub repository: https://github.com/lzillich/VS_multi_omics_cocaine, https://doi.org/10.5281/zenodo.14637857.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

4.10 Acknowledgements

Funding supporting this study was provided by the German Federal Ministry of Education and Research (BMBF) within the e:Med research program SysMedSUDs: "A systems-medicine approach toward distinct and shared resilience and pathological mechanisms of substance use disorders" (01ZX01909 to R.S., P.K., M.R., A.C.H., and S.H.W.). Additionally, funding was provided by the Deutsche Forschungsgemeinschaft (DFG) through the collaborative research center TRR265: "Losing and Regaining Control over Drug Intake" (Heinz et al., 2020; Spanagel et al., 2024) (project ID 402170461 to S.H.W., R.S., A.C.H., and M.R.), the Hetzler Foundation for Addiction Research (to A.C.H.), and the ERA-NET program: Psi-Alc (01EZ1908). The project has been carried out using the Mannheim (CIMH) infrastructure of the German Center for Mental Health (DZPG). We thank the Proteomics Core Facility at the European Molecular Biology Laboratory (EMBL, Heidelberg, Germany) for their support in proteomic analyses. We further thank Elisabeth Röbel and Claudia Schäfer-Arnold for their technical assistance. The graphical abstract and Figures 4.1A and 4.1B were created using BioRender (https://www.biorender.com).

4.11 Author contributions

Conceptualization, E.Z., L.Z., M.R., R.S., and S.H.W.; methodology, E.Z., A.A., M.M.N., A.C.R., C.C.W., J.F., and L.Z.; resources, G.T., N.M., P.K., and A.C.H.; data curation, E.Z., A.A., A.C.R., D.A., H.B., and L.Z.; data analysis, E.Z., D.A., H.B., and L.Z.; investigation, E.Z., A.A., A.C.R., D.A., H.B., A.C.H., M.R., R.S., S.H.W., and L.Z.; writing – original draft, E.Z.; writing – review & editing, E.Z., A.A., A.C.R., D.A., H.B., J.F., N.M., G.T., M.M.N., A.C.H., C.C.W., M.R., P.K., R.S., S.H.W., and L.Z.; supervision, J.F., M.R., C.C.W., P.K., R.S., S.H.W., and L.Z.; project administration, M.M.N., P.K., R.S., A.C.H., M.R., and S.H.W.; funding acquisition, P.K., R.S., A.C.H., M.R., and S.H.W.

4.12 Declaration of interests

The authors declare that there are no competing interests.

4.13 STAR★Methods

4.13.1 Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human postmortem brain tissue of the ventral striatum	Douglas Bell Canada Brain Bank (DBCBB), Montreal, Canada	https://douglas.research.mcgill.ca/douglas-bell-canada-brain-bank/
Chemicals, peptides, and recombinant proteins		

cOmplete™ Mini protease inhibitor	Sigma Aldrich	Cat# 11836170001
Pierce® RIPA buffer	Thermo Fisher Scientific	Cat# 89900
NuPAGE™ 4x LDS Sample Buffer	Thermo Fisher Scientific	Cat# NP0007
NuPAGE™ 10x Reducing Agent	Thermo Fisher Scientific	Cat# NP0009
Dithiothreitol	Biomol	Cat# 04010.25
2-chloroacetamide	Sigma-Aldrich	Cat# C0267
HEPES	Sigma-Aldrich	Cat# H23830
Trypsin (sequencing grade)	Promega	Cat# V5111
Acetonitrile	Chemsolute	Cat# 2697
Hydroxylamine	Sigma-Aldrich	Cat# 438227
Ammonium formate	Sigma-Aldrich	Cat# 78314
Formic acid	Fisher Chemical	Cat# A117-50
Trifluoroacetic acid	Sigma-Aldrich	Cat# 80457
DMSO	Sigma-Aldrich	Cat# 276855
Critical commercial assays		
miRNeasy Tissue/Cells Advanced Micro Kit	Qiagen	Cat# 217684
NEXFLEX small RNA-seq Kit v3	Perkin Elmer	Cat# NOVA-5132-06
NEBNext Ultra II Directional RNA Library Prep Kit	New England Biolabs	Cat# E7760L
10X Genomics Chromium Nuclei Isolation Kit with RNase Inhibitor	10X Genomics	Cat# PN-1000494
Dual Index Kit TT Set A	10X Genomics	Cat# PN-1000215
Chromium Next GEM Chip G Single Cell Kit	10X Genomics	Cat# PN-1000127
10X Genomics Chromium Next GEM Single Cell 3′ Kit v3.1 (Dual Index)	10X Genomics	Cat# PN-1000268
TMT10plex™ Isobaric Label Reagent	Thermo Fisher Scientific	Cat# 90110
TMTpro™ 16plex Label Reagent Set	Thermo Fisher Scientific	Cat# A44520
Deposited data		
Raw sequencing and proteomics datasets	This paper	EGA: https://ega-archive.org/studies/EGAS50000000623 EGA: https://ega-archive.org/studies/EGAS00001007945
Pre-processed snRNA-seq data from nucleus accumbens of adult male rats undergoing a repeated cocaine exposure paradigm	Phillips et al. (2023)	NCBI GEO: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE222418
Software and algorithms		
Cell Ranger, v.7.1.0	10x Genomics	https://www.10xgenomics.com/support/software/cell-ranger/latest
R, v.4.2.1	R Core Team (2021)	https://www.r-project.org
FastQC, v.0.12.1	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

TrimGalore, v.0.6.10	Babraham Bioinformatics	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
STAR, v. 2.7.10b	Dobin et al. (2013)	https://github.com/alexdobin/STAR
Rsubread, v.2.12.3	Liao et al. (2019)	https://bioconductor.org/packages/release/bioc/html/Rsubread.html
Salmon, v.1.10.0	Patro et al. (2017)	https://combine-lab.github.io/salmon/
tximport, v.1.26.1	Soneson et al. (2015)	https://bioconductor.org/packages/release/bioc/html/tximport.html
Seurat, v.5.0.1	Hao et al. (2024)	https://satijalab.org/seurat/
SoupX, v.1.6.2	Young and Behjati (2020)	https://github.com/constantAmateur/SoupX
Fragpipe, v.20.0	Kong et al. (2017)	https://fragpipe.nesvilab.org
MSFragger, v.3.8	Kong et al. (2017)	https://msfragger.nesvilab.org
limma, v.3.54.2	Ritchie et al. (2015)	https://bioconductor.org/packages/release/bioc/html/limma.html
vsn, v.3.66.0	Huber et al. (2002)	https://www.bioconductor.org/packages/release/bioc/html/vsn.html
DESeq2, v.1.38.3	Love et al. (2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
clusterProfiler, v.4.6.2	Yu et al. (2012)	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
enrichplot, v.1.18.3	Yu (2024)	https://www.bioconductor.org/packages/release/bioc/html/enrichplot.html
ComplexHeatmap, v.2.14.0	Gu et al. (2016)	https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html
UpSetR, v.1.4.0	Conway et al. (2017)	https://github.com/hms-dbmi/UpSetR
CIBERSORT, v.1.04	Newman et al. (2015)	https://cibersortx.stanford.edu
BayesianFirstAid, v.0.1	Bååth (2014)	https://github.com/rasmusab/bayesian_first_aid

GeneOverlap, v.1.34.0	Shen L (2023)	https://bioconductor.org/packages/release/bioc/html/GeneOverlap.html
WGCNA, v.1.72.1	Langfelder and Horvath (2008)	https://cran.r-project.org/web/packages/WGCNA/index.html
MOFA2, v.1.8.0	Argelaguet et al. (2018)	https://www.bioconductor.org/packages/release/bioc/html/MOFA2.html
jyluMisc, v.0.1.5	GitHub repository	https://github.com/lujunyan1118/jyluMisc
hdWGCNA, v.0.2.26	Morabito et al. (2023)	https://smorabit.github.io/hdWGCNA/
CellChat, v.2.1.2	Jin et al. (2021)	https://github.com/jinworks/CellChat
Other		
Resource website for data analysis of this paper	This paper	https://github.com/lzillich/VS_multi_omics_cocaine, Zenodo: https://doi.org/10.5281/zenodo.14637857

4.13.2 Experimental model and study participant details

4.13.2.1 Human postmortem brain tissue

Postmortem human brain tissue of the ventral striatum from $n = 20$ individuals with cocaine use disorder and $n = 21$ unaffected individuals was obtained from the Douglas Bell Canada Brain Bank (DBCBB) in Montreal, Canada. Ventral striatum was dissected in brain sections equivalent to plate 15 (−7.5 mm from the center of the anterior commissure) of the human brain atlas by Mai et al. (2007). Tissue was taken rostral to the anterior commissure and ventral to the tip of the anterior limb of internal capsule covering central, medial, and lateral parts of the nucleus accumbens as well as caudate and putamen fundus regions. Tissue sampling and acquisition of phenotype information at DBCBB was performed based on their established ethical guidelines. Our multi-omics study was approved by the Ethics Committee II of the University of Heidelberg, Medical Faculty Mannheim, Germany, under the register number 2021-681. Inclusion criteria were donor age >18 and a DSM-IV diagnosis of cocaine dependence. Instead of cocaine dependence, throughout this study, the more recent terminology from DSM-V i.e., cocaine use disorder is used. Exclusion criteria were other substance use disorders except alcohol use disorder and a diagnosis of severe neurodevelopmental or (neuro-)psychiatric disorders except major depressive disorder (MDD). To address strong inter-individual heterogeneity of human individuals and its effects on statistical analyses, we aimed for a homogeneous sample in our study resulting in a final

cohort of n = 41 male individuals with European American ancestry consisting of n = 20 individuals with CUD and n = 21 non-affected individuals assigned to the Ctrl group. Phenotype information other than CUD status include tissue donor age, sex (reported as biological sex), ethnic background, postmortem interval (PMI), brain pH value, additional psychiatric diagnoses, toxicology at death and detailed cause and manner of death. A detailed description of phenotypes is provided in Table S4.1. Prior to all postmortem brain sample processing steps during omics data generation, randomization based on CUD status was performed to minimize confounding technical and batch effects.

4.13.3 Method details

4.13.3.1 miRNA sequencing

Extraction of miRNAs from was performed using the miRNeasy Tissue/Cells Advanced Micro Kit (Qiagen, Hilden, Germany) resulting in a total RNA preparation that contains small RNAs such as miRNAs as well as long RNAs. For all postmortem brain samples, a total of 5mg tissue was used during total RNA extraction. RNA quality of the n = 41 total RNA preparations was assessed on an Agilent TapeStation 4200 (Agilent, Santa Clara, CA, USA) and n = 40 samples remained based on an RNA integrity number cut-off of >5.5. Library preparation for miRNA-sequencing was performed using the NEXFLEX small RNA-seq kit v3 (PerkinElmer, Waltham, MA, USA). Small-RNA libraries were sequenced on an Illumina NovaSeq 6000 (Illumina, San Diego, CA, US) with an average of 10 million reads per sample (1x50bp).

4.13.3.2 RNA sequencing

RNA was extracted as total RNA from 5mg of postmortem human brain tissue using the miRNeasy Tissue/Cells Advanced Micro Kit (Qiagen, Hilden, Germany). RIN values were determined as previously described in the miRNA extraction procedure with RIN>5.5 as the RNA quality cut-off resulting in n = 40 samples for which library preparation was performed. Transcriptomic profiling of the VS samples was based on an NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) that was used for library preparation with rRNA depletion. RNA sequencing for the n = 40 samples was performed on an Illumina NovaSeq 6000 device with an average of 60 million read pairs (2x100bp).

4.13.3.3 Single-nuclei RNA sequencing

Nuclei were isolated from n = 8 CUD cases and n = 8 unaffected Ctrl individuals. Samples for snRNA-seq were selected from the full cohort of n = 41 individuals by the amount of available tissue to perform all preparations (miRNA, RNA, protein,

and snRNA) from the same tissue sample. Samples with low postmortem interval and higher RIN value as measured in the bulk RNA-seq analysis were preferred. For the nuclei isolation, 10mg of frozen postmortem tissue was used to minimize the amount of free-floating debris. Samples were processed as indicated in the 10X Genomics Chromium Nuclei Isolation kit with RNase Inhibitor manufacturers protocol (10x Genomics, Pleasanton, CA, USA). Briefly, the tissue was dissociated in lysis buffer, filtered, and the cellular debris were removed. After multiple centrifugation and washing steps, a clean nuclei suspension was obtained. Nuclei were automatically counted on LUNA-FL Dual Fluorescence Cell Counter while excluding particles smaller than 5µm and bigger than 15µm of diameter from the counting. Single-nuclei RNA-seq (snRNA-seq) libraries were generated using the 10X Genomics Chromium Next GEM Single Cell 3′ Kit v3.1 (Dual Index). A total of 10,000 nuclei per sample were loaded into chromium chips. Libraries were prepared following the provided 10X Genomics Dual Index Kit protocol and sequenced using an Illumina NovaSeq6000 device with 50,000 read pairs per cell.

4.13.3.4 TMT quantitative proteomics

Proteomic profiling was based on a Radioimmunoprecipitation Assay (RIPA) buffer lysate of postmortem human brain tissue samples. For each sample preparation, 50-70mg postmortem brain tissue was mechanically disrupted and homogenized under a liquid nitrogen atmosphere. Disrupted tissue was transferred to a reaction tube and was suspended in lysis buffer Pierce RIPA (Thermo Fisher Scientific, Waltham, MA, USA) with protease inhibitor (cOmplete Mini, Roche, Basel, Switzerland). After a lysis period of 30 min at 4°C, the suspension was centrifuged for 10 min at 12,500xg (4°C) and the supernatant was kept for further lysate preparation steps. Protein concentration was measured for each sample on Direct Detect Assay-free membrane cards using a Direct Detect Spectrometer (both Merck Millipore, Burlington, MA, USA). RIPA lysate containing 50µg total protein was processed for SDS gel electrophoresis. NuPAGE 4x LDS Sample Buffer and NuPAGE 10x Reducing Agent (both Thermo Fisher Scientific, Waltham, MA, USA) were added to the protein lysate and incubated for 10 min at 75°C. From this sample preparation, a total of 25µg protein were analyzed in TMT proteomics.
The reduction of disulfide bonds on cysteine was conducted using dithiothreitol (56°C, 30 min, 10 mM in 50 mM HEPES, pH 8.5) followed by alkylation with 2-chloroacetamide (room temperature (RT), in the dark, 30 min, 20 mM in 50 mM HEPES, pH 8.5). The SP3 protocol (Hughes et al., 2014; Hughes et al., 2019) was employed for sample clean-up, and trypsin (sequencing grade, Promega, Madison, WI, USA) was added at an enzyme-to-protein ratio of 1:50 for overnight digestion at 37°C (in 50 mM HEPES). The n = 40 samples were randomized into three TMT multiplex batches (two with 16 samples each and one with 10 samples) based on CUD status and phenotypic covariates. Peptides were labelled with either TMT10plex (n = 8 sample batch) Werner et al. (2014) or TMT16plex (2x n = 16

sample batches) Thompson et al. (2019) Isobaric Label Reagent (Thermo Fisher Scientific, Waltham, MA, USA) according the manufacturer's instructions. In brief, 0.8 mg of reagent was dissolved in 42 mL of acetonitrile (100%) and 8 µL was added, followed by an incubation period of 1 h at room temperature. The samples were incubated with 8 µL of 5% hydroxylamine for 15 min at room temperature. The samples from a given measurement batch were combined and desalted on an OASIS HLB µElution Plate (Waters, Milford, MA, USA). The offline high pH reverse phase fractionation was conducted on an Agilent 1200 Infinity high-performance liquid chromatography system, equipped with a Gemini C18 column (3 µm, 110 Å, 100 x A 1.0 mm Phenomenex Torrance, CA, USA column was installed with a Gemini C18 4 × 2.0 mm SecurityGuard cartridge (Phenomenex, Torrance, CA, USA) as a guard column. The binary solvent system comprised 20 mM ammonium formate (pH 10.0) (A) and 100% acetonitrile as the mobile phase (B). The flow rate was set to 0.1 mL/min. Peptides were separated using a gradient of 100% A for 2 min, to 35% B in 59 min, to 85% B in another 1 min and kept at 85% B for an additional 15 min, before returning to 100% A and re-equilibration for 13 min. A total of 48 fractions were collected which were subsequently pooled into 12 fractions. Pooled fractions were dried under vacuum centrifugation, reconstituted in 10 µL 1% formic acid, 4% acetonitrile and then stored at −80°C until LC-MS analysis.

An UltiMate 3000 RSLC nano LC system (Dionex, Sunnyvale, CA, USA) fitted with a trapping cartridge (µ-Precolumn C18 PepMap 100, 5µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (nanoEase M/Z HSS T3 column 75 µm × 250 mm C18, 1.8 µm, 100 Å, Waters, Milford, MA, USA) was coupled to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) using the Nanospray Flex ion source in positive ion mode. The samples were applied to the trapping column at a constant flow rate of 30 µL/min with 0.05% trifluoroacetic acid in water for a period of 4 min. After switching in line with the analytical column peptides were eluted at a constant flow of 0.3 µL/min using the method described below. The binary solvent system comprised 0.1% formic acid in water with 3% DMSO (solvent A) and 0.1% formic acid in acetonitrile with 3% DMSO (solvent B). The percentage of solvent B was increased from 2% to 8% in 4 min (n = 10 TMT batch) or in 2 min (n = 16 TMT batches), from 8% to 28% in 72 min (n = 10 TMT batch) or 104 min (n = 16 TMT batches), to 38% (n = 10 TMT batch) or to 40% (n = 16 TMT batches) in another 4 min and finally to 80% B in 4 min, followed by re-equilibration back to 2% B in 4 min. The peptides were introduced into the mass spectrometer via a Pico-Tip Emitter 360 µm OD x 20 µm ID; 10 µm tip (New Objective, Woburn, MA, USA) and an applied spray voltage of 2.4 kV. The capillary temperature was set to 275°C. A full mass scan was conducted with a mass range of 375–1500 m/z in profile mode in the orbitrap with a resolution of 120000. The maximum filling time was set at 50 ms, with a limit of $4x10^5$ ions. The data-dependent acquisition (DDA) was conducted with the Orbitrap resolution set to 30000, with a fill time of 94 ms,

and a limitation of $1x10^5$ ions. A normalized collision energy of 36 (n = 10 TMT batch) or 34 (n = 16 TMT batches) was applied. MS^2 data was acquired in profile mode. The define first mass was set to 110 m/z.

4.13.4 Quantification and statistical analysis

All statistical analyses were performed in the statistical computing environment R, version 4.2.1. If not otherwise stated, Benjamini-Hochberg (FDR) (Benjamini & Hochberg, 1995) correction was performed to adjust for multiple testing.

4.13.4.1 Sequencing data processing

Raw sequencing data from miRNA-sequencing were processed using FastQC v.0.12.1 and quality metrics were inspected. Sequencing adapters were trimmed using TrimGalore v.0.6.10. with automatic adapter detection and a minimum read length threshold of 18bp. Success of adapter trimming was evaluated in a post-processing run of FastQC. Alignment of miRNA-sequencing reads to the human reference genome (hg38) was performed using STAR v. 2.7.10b (Dobin et al., 2013). First, a genome index for the GENCODE GRCh38 primary assembly reference genome together with the GENCODE v.43 genome annotation file was created using STAR with --sjdbOverhang 49 for miRNA-sequencing data. Next, alignment was performed for the n = 40 fastq-files while filtering for multi-mapping using the --outFilterMultimapNmax 20 flag. Quantification of miRNAs was performed using featureCounts as implemented in Rsubread v.2.12.3 (Liao et al., 2019) with the miRNA-specific hsa.gff3 reference file from miRBase (https://www.mirbase.org/download/, release 22.1).
For the RNA-seq dataset, raw fastq-files were inspected in FastQC as described for the miRNA-sequencing dataset. The GRCh38 reference genome was indexed in STAR using the --sjdbOverhang 100 flag. Alignment in STAR was performed using default parameters. For the quantification of transcripts, featureCounts was used with the GENCODE v.43 genome gtf-file as the reference annotation.
Raw snRNA-seq data from the n = 16 VS tissue samples was processed using Cell Ranger v.7.1.0 (10x Genomics) (Zheng et al., 2017) and resulting feature-barcode were imported in Seurat v.5.0.1 (Hao et al., 2024). Low quality nuclei were removed from the analysis using the following QC parameters: 900 < nFeature_RNA <8500 and mitochondrial gene fraction <10%. After QC, ambient RNA correction was performed using SoupX v.1.6.2 (Young & Behjati, 2020). Expression data matrices from the n = 16 samples were merged into a single Seurat object. Count normalization was performed using “NormalizeData”, followed by the identification of variable features (“FindVariableFeatures”) and data scaling (“ScaleData”). Next, “RunPCA” and “RunUMAP” were applied for dimensionality reduction and the n = 16 individual datasets were integrated using “IntegrateLayers” with “method = HarmonyIntegration”. “FindClusters” was applied with a resolution parameter of 0.1. The association of clusters with

technical parameters such as number of features and mitochondrial gene counts was inspected leading to the removal of one cluster that was strongly enriched for mitochondrial genes. Re-clustering at a resolution of 0.12 resulted in the final object with n = 12 cell type clusters based on expression data from n = 20,759 nuclei.

Proteomics data preprocessing and analysis was performed based on an adapted version of an analysis pipeline by Frank Stein (EMBL, Heidelberg, Germany). Preprocessing of the proteome data was performed using Fragpipe v20.0 (MSFragger v.3.8) (Kong et al., 2017) by searching against a *Homo sapiens* proteome database (UP000005640, October 2022, 20594 entries) plus common contaminants and reversed sequences. The following modifications were included into the search parameters: Carbamidomethyl on Cysteine and TMT10/16 on lysine as fixed modifications, protein N-term acetylation, oxidation on methionine and TMT10/16 on N-termini as variable modifications. A mass error tolerance of 20 ppm was applied to both precursor and fragment ions. Trypsin was set as protease with a maximum of two missed cleavages. The minimum peptide length was set to seven amino acids. At least two unique peptides were required for protein identification. The false discovery rate on peptide and protein level was set to 0.01. The raw reporter intensities from the three TMT plexes were extracted from the raw tsv output files of FragPipe. Contaminants were removed, and only proteins quantified with at least two Razor peptides were included in the analyses. Due to the intrinsic nature of postmortem human brain studies characterized by strong interindividual heterogeneity, we did not perform imputation of missing values but restricted the proteomic analysis to proteins that were identified in all of the three TMT experiments. A total of n = 4270 proteins passed the quality control filters. Log2-transformed raw TMT reporter ion intensities were first cleaned for batch effects using the "removeBatchEffects" function from limma v.3.54.2 (Ritchie et al., 2015) to minimize the influence of TMT batch on the intensity values. Next, batch corrected data was variance stabilization normalized using the vsn package v.3.66.0 (Huber et al., 2002).

4.13.4.2 Quality control of bulk-level datasets

To identify potential outliers, we performed sample clustering analysis in the miRNA-seq, RNA-seq, and proteomics datasets individually using principal component analysis (PCA) on variance stabilization transformed data. Here, we found two samples (CUD_3 and CUD_9) that consistently separated from the remaining samples across bulk-level datasets (Figure S4.1B). In a cell type deconvolution analysis on bulk RNA expression signatures using CIBERSORT (Figures S4.1C and S4.1D), we found that for these samples only a minimal proportion of medium spiny neurons (MSNs), the major neuronal cell type in the VS, was estimated. To minimize the influence of differential cell type proportion on the results of bulk-level analyses, we excluded the two samples predicted to lack

a major cell type of the VS. We thus restricted differential expression and downstream analyses including WGCNA and MOFA to the remaining n = 38 samples.

4.13.4.3 Differential microRNA expression analysis

Differential expression (DE) testing for the miRNA expression dataset was performed in DESeq2 v.1.38.3 (Love et al., 2014) with covariate adjustment for donor age, PMI, brain pH, and RIN using the following DESeq2 model: *Expr(miRNA) ~ CUD + age + PMI + pH + RIN*. Threshold criteria for differential miRNA expression in the CUD vs. Ctrl comparison were absolute log2FC larger than 0.07 (5% change in miRNA expression), an association with $p < 0.05$ for nominal significance, and FDR-adjusted q < 0.05 for significance after multiple testing correction.

4.13.4.4 Differential RNA expression analysis

Similar to the miRNA-sequencing dataset, DE testing on the raw count matrix in DESeq2 was performed using the same set of covariates in the statistical model: *Expr(RNA) ~ CUD + age + PMI + pH + RIN*. Again, criteria for differential expression were absolute log2FC > 0.07, $p < 0.05$ for nominal significance, and FDR-adjusted q < 0.05 for transcriptome-wide significance. To evaluate functional enrichment of DE results, a Gene Ontology (GO) enrichment analysis was performed based on transcripts with FDR<0.25 to obtain a lenient but still statistically stringent set of genes. The GO enrichment analysis was performed using the compareCluster function of the R package clusterProfiler v.4.6.2 (Yu et al., 2012). Terms with significant enrichment (FDR-adjusted q < 0.05) were extracted from the GO enrichment results. For visualization, functional modules were generated using the emapplot function of enrichplot v.1.18.3.

4.13.4.5 Differential expression analysis of snRNA-seq data

Annotation of clusters in the VS snRNA dataset was based on the expression of marker genes that have been previously used in other snRNA-seq studies of the brain (Brenner et al., 2020; Tran et al., 2021). DE testing was restricted to the n = 7 major cell type clusters consisting of at least 100 cells from each condition (CUD/Ctrl). While negative binomial mixed models or pseudo-bulk approaches are suitable analysis tools that consider the hierarchical nature of single-cell datasets, a relatively large sample size is required to obtain sufficient statistical power, particularly in human datasets characterized by strong inter-individual heterogeneity. To perform the DE analysis using a sufficiently powered statistical approach in our sub-sample of n = 16 individuals, we performed a CUD vs. Ctrl comparison using the Wilcoxon rank-sum test in "FindMarkers" on robustly expressed genes (min.pct = 0.25) in each cluster and then applied strict filtering

criteria to the Bonferroni-significant results to select only the top relevant DEGs for further characterization in downstream analyses. DEGs were prioritized based on strong effect size (|log2FC| > 0.5) and significance filtering was performed for p.adjust<0.001 to limit the number of potential false positive associations. For visualization of DE results, a DE gene heatmap was generated using ComplexHeatmap v.2.14.0 (Gu et al., 2016) and the most strongly deregulated CUD-associated DEGs based on log2FC from each cluster were highlighted by gene name. Functional enrichment of cell type specific DEGs within biological pathways was evaluated using GO enrichment analysis. Significant results (FDR-adjusted q < 0.05) were visualized in an enrichment map. DE gene overlap across evaluated clusters was analyzed in an upset plot created using UpSetR v.1.4.0 (Conway et al., 2017).

4.13.4.6 Differential protein expression analysis

Differential expression testing for proteins in CUD cases and control individuals was performed based on the normalized, log2-transformed reporter intensities. The following statistical model was used in the model matrix that was given as an argument to the 'lmFit' function of limma: *Expr(protein) ~ CUD + age + PMI + pH + batch (TMT)*. FDR-correction was used to adjust *p*-values for multiple testing. Protein DE cut-offs were the same as used in the miRNA-seq and RNA-seq analysis: |log2FC| > 0.07, $p < 0.05$ for nominal significance, FDR-adjusted q < 0.05 for proteome-wide significance. Using the same filtering criteria as in the GO enrichment analysis with DEGs, proteins associated with CUD at FDR<0.25 were extracted from the results and investigated in a GO enrichment analysis in clusterProfiler.

4.13.4.7 Cell type deconvolution analysis

Differences in the distribution of brain cell types across samples might lead to differential RNA expression independent of other phenotypes. We thus performed a cell type deconvolution analysis using CIBERSORT (Newman et al., 2015) to evaluate the relationship between cell type percentage and differential expression results in the RNA-seq data. As we have generated single-nuclei RNA-seq data from a subset of the VS postmortem brain samples in this study, we created a customized gene expression reference matrix for the ventral striatum based on our snRNA data from $n = 16$ individuals. From the snRNA dataset, cell type specific expression data was extracted for glial cells such as astrocytes, oligodendrocytes, oligodendrocyte precursor cells, and microglia. To obtain robust estimates, the different neuronal clusters were summarized to a medium-spiny neuron (MSN) cluster containing D1-and D2-MSNs (Inh_MSN) and an inhibitory GABA cluster (Inh_GABA) containing all other inhibitory striatal neurons. Next, from the transcriptome-wide gene x cell normalized count matrix, we selected potential marker genes characterized by at least 10-times stronger expression in one cell

type compared to all other cell types. The filtered gene expression matrix (3,081 genes in 20,492 cells) was used as the input dataset for the CIBERSORT.jar distribution resulting in a customized reference matrix for the VS. Using the CIBERSORT.R script (v.1.04, https://cibersortx.stanford.edu), CIBERSORT cell type estimates were generated for the bulk RNA-seq data in n = 40 samples using normalized counts from DESeq2 (Figure S4.1C). To evaluate the accuracy of cell type estimation based on our customized reference matrix for the VS, we compared the measured cell type proportions as determined in snRNA-seq with the CIBERSORT estimates for the n = 16 samples for which both bulk and snRNA-seq data are available. Pearson correlation coefficients ranged from r = 0.44 to r = 0.99 (median r = 0.86) confirming successful generation of a customized reference matrix for the estimation of cell types in the VS (Figure S4.1D). Using the snRNA-seq derived reference matrix, differences in cell type proportion between CUD cases and control individuals were tested using the Bayesian estimation procedure from BEST (Kruschke, 2013) (R package BayesianFirstAid v.0.1).

4.13.4.8 Transcriptome-proteome correlation analysis

As few is known about the correlation between RNA and protein levels in postmortem human brain, the overall relationship between transcriptome and proteome was investigated using a correlation analysis approach. Following the protocol from Yang and Gorski (2022) we performed preprocessing of transcriptomic and proteomic datasets. To make gene expression levels compatible with protein levels, the RNA-seq data was normalized to gene length and sequencing depth, and transcript per million (TPM) values were generated. For this, we performed pseudo-alignment of RNA-seq raw data to the GRCh38 primary assembly reference transcriptome using Salmon v.1.10.0 (Patro et al., 2017). From the Salmon quantification files, TPM estimates were extracted using tximport v.1.26.1 (Soneson et al., 2015) followed by log2-transformation of TPM values. For the proteomic dataset we used the batch corrected, log2-transformed and vsn-normalized TMT intensity data. We performed filtering of RNA-seq and proteomics data for individuals that have both gene expression and protein data (n = 39) and from the RNA-seq dataset we kept only protein coding genes that were detected in the proteomics dataset (n = 3,935). At the sample level, within-individual correlation of transcriptome and proteome was determined using Pearson correlation. Next, we averaged RNA and protein expression levels across samples and calculated the correlation between mean RNA mean protein expression values for the n = 3,935 genes. Further, Pearson correlation was assessed in a gene-centered approach resulting in a distribution of n = 3,935 correlation coefficients. Based on the ranking of genes by correlation coefficients, we performed pre-ranked gene set enrichment analysis (GSEA) as implemented in clusterProfiler (Yu et al., 2012) to evaluate functional enrichment of strongly positively and negatively correlated genes within biological pathways. Finally, we investigated RNA-protein

correlation for each gene separately in CUD and Ctrl samples to identify genes that show the strongest difference in RNA-protein correlation associated with CUD/Ctrl status. A difference score ΔR was calculated by subtracting RNA-protein correlation coefficients in Ctrl from RNA-protein correlation coefficients in CUD samples (R_{CUD}-R_{Ctrl}).

4.13.4.9 WGCNA

To evaluate co-expression patterns in the miRNA-seq (n = 1,542 miRNAs), RNA-seq (n = 22,685 RNAs), and proteomic datasets (n = 4,270 proteins), a weighted correlation network analysis (WGCNA, package v.1.72.1) (Langfelder & Horvath, 2008) was performed. Using high-dimensional input datasets, WGCNA applies a pairwise correlation and hierarchical clustering approach to identify co-expression modules that are then related to trait variables such as the CUD phenotype. Normalized and variance stabilization transformed counts/TMT reporter intensities were used as the input datasets. Soft-power thresholds to achieve the criterion of scale-free topology ($R^2 > 0.9$) were determined for each dataset individually by running the pickSoftThreshold function resulting in powers of 6, 7, and 9 for the miRNA-seq, RNA-seq and proteomics dataset, respectively. Next, automated network construction was performed with the parameters minModuleSize = 10 for miRNA data, minModuleSize = 20 for RNA and protein datasets, mergeCutHeight = 0.15, and maxBlockSize = 36,000. Pearson correlation coefficients of the module eigengene, corresponding to the first principal component of the module's expression matrix, with trait data including CUD and other phenotypes such as pH, RIN, PMI, and CIBERSORT estimated cell type proportions, were determined in each dataset. We further performed a linear regression analysis for CUD status on module eigengenes where we performed adjustment for covariates age, pH, PMI, RIN, and proteomics batch using the linear models ME ~ CUD + age + pH + PMI + RIN (RNA) and ME ~ CUD + age + pH + PMI + batch (protein). Co-expression modules characterized by i) significant correlation with CUD ($p < 0.05$) and ii) significant association with CUD adjusted for covariates were characterized using GO enrichment analysis as implemented in the enrichGO function of clusterProfiler. Further, to evaluate the relationship between modules across datasets, module eigengene correlation was determined between CUD-associated modules in RNA and protein data using Pearson correlation. Finally, we performed an overlap analysis using a Fisher Test as implemented in GeneOverlap (Shen L, 2023) to evaluate potential enrichment of differentially expressed genes and proteins in CUD-associated co-expression modules.

4.13.4.10 Multi-omics factor analysis

An integrative multi-omic analysis of miRNA-seq, RNA-seq, and proteomic datasets was performed using Multi Omics Factor Analysis (MOFA) (Argelaguet

et al., 2018). MOFA as implemented in the R package MOFA2 v.1.8.0 provides a statistical framework for high-dimensional (omics) data integration leveraging factor analysis for the unsupervised identification of lower-dimensional factor representations of the input datasets. Relationship of factors to trait variables such as CUD is assessed in downstream analyses allowing the identification of features (i.e., miRNAs, transcripts, or proteins) that show CUD-associated variability. During data preprocessing, normalized and variance stabilization transformed counts/TMT intensities were z-scaled using the mscale function from jyluMisc v.0.1.5 and the resulting matrices were used as the input datasets in MOFA. Due to the different dimensionalities of the omics datasets and its potential negative influence on MOFA model performance (n = 1,542 miRNAs, n = 22,685 RNAs, and n = 4,270 proteins), the top 4,270 highly variable genes from RNA-seq data were filtered and thereby matched to the size of the proteomic dataset. Expression data for all n = 1,542 miRNAs was included. In MOFA, default data and model options were used, whereas training options were modified using "convergence mode" – slow and a "drop_factor_threshold" of 0.01 to drop factors from the model than explain less than 1% variance in each view. The learned factors were inspected and factors significantly associated with CUD were investigated using GSEA based on the ranking of RNA and protein features by factor weights. GSEA was performed using the gseGO function of clusterProfiler.

4.13.4.11 hdWGCNA

Cell type specific co-expression signatures were investigated using high-dimensional WGCNA (Morabito et al., 2023) (hdWGCNA, R package v.0.2.26, documentation from https://smorabit.github.io/hdWGCNA/index.html). hdWGCNA was performed based on the code implementation from its source publication (https://github.com/smorabit/hdWGCNA_paper). To ensure sufficient cluster sizes for robust module detection, D1-MSN and D2-MSN clusters were combined to an "Inh_MSN" cluster and inhibitory neuron clusters GABAergic-1, GABAergic-2, and GABAergic-3 were condensed to an "Inh_GABA" cluster.
Iteratively, hdWGCNA was performed in each cell type resulting in co-expression module labels according to cell types. Metacells were generated with nearest-neighbour k = 25 followed by automated network construction using "TestSoftPowers" and "ConstructNetwork" functions with default parameters. Resulting co-expression modules for each cell type cluster were inspected for cluster-specific module eigengene expression using "ModuleFeaturePlot". A co-expression module was considered cell type-specific if the module eigengene showed strongest expression in the cell type with the same name as the co-expression module and the module eigengene expression pattern was robust across the respective cell type cluster i.e., the association is not driven by a few cells only. Next, differential module eigengene testing was performed to identify co-expression modules that show significant differences

(upregulation/downregulation) between CUD cases and Ctrl individuals. For further downstream analyses, we prioritized modules that i) showed strongest expression in the cell type cluster that was used for constructing the co-expression networks thereby addressing cell type specificity of modules and ii) displayed statistically significant module eigengene differences associated with CUD ($p < 0.05$) indicating differential abundance of co-expression patterns in CUD (Table S4.6A). In the resulting module subset, pathway enrichment analyses were performed using the "RunEnrichr" function in hdWGCNA using GO and KEGG databases as reference. Results were ranked by the combined.score metric from Enrichr defined as log(p) from Fisher-Test multiplied by the Z score as the deviation from the expected rank. Using GeneOverlap, we tested the enrichment of cluster-specific CUD-associated DE genes within co-expression modules.

4.13.4.12 Consensus hdWGCNA

As an integrative approach, we performed a cross-species consensus network analysis in hdWGCNA based on publicly available snRNA-seq data from a repeated cocaine-exposure model in rats (Savell et al., 2020). Sequencing data for the nucleus accumbens from male rats was downloaded from GEO (accession number: GSE222418) and processed using the same analysis pipeline as described for the analysis of the human dataset. Clusters in the rat dataset were annotated based on cell type markers from the original publication and neuronal subclusters were condensed into Inh_MSN and Inh_GABA clusters as previously described for the human dataset. Network construction was performed on expression data for homologous genes across species. From the identified consensus modules, we selected modules that showed i) strongest module gene expression in the cell type of interest and ii) significant differential module eigengene association with CUD (Table S4.6D). The overlap of module genes between CUD-associated DME modules from the analysis in human CUD and the consensus co-expression modules across species was evaluated using GeneOverlap. Finally, pathway enrichment analysis was performed in the CUD-associated consensus modules that most strongly overlapped between the human and consensus analyses.

4.13.4.13 CellChat

We used CellChat (Jin et al., 2021) (R package v.2.1.2) to analyze CUD-associated expression changes of ligand-receptor pairs in MSNs and astrocytes. Following the documentation for combined analysis of multiple datasets (https://github.com/jinworks/CellChat/blob/main/tutorial/Comparison_analysis_of_multiple_datasets.html), construction of CellChat objects was performed individually in CUD and Ctrl nuclei using the human reference dataset for ligand-receptor interaction (CellChatDB.human). Next, the individual objects were merged using "mergeCellChat". A differential expression analysis for ligand receptor (LR) pairs was performed using "identifyOverExpressedGenes" and

"netMappingDEG" to identify ligands and receptors with statistically significant deregulation in CUD using tresh.pc = 0.1, thresh.fc = 0, and thresh.p = 0.05 cut-offs. Significant up- and downregulated LR pairs were filtered for interactions with a minimum of 5% change in ligand expression ($|\log2FC > 0.07|$) and an expression change of its receptor into the same direction. The "netVisual_chord_gene" function was used to visualize differential LR interactions individually for up- and downregulated LR pairs in CUD. A donut plot was generated to summarize the distribution of signaling type annotations among up- and downregulated LR pairs.

4.14 Supplemental information

Supplementary Figures

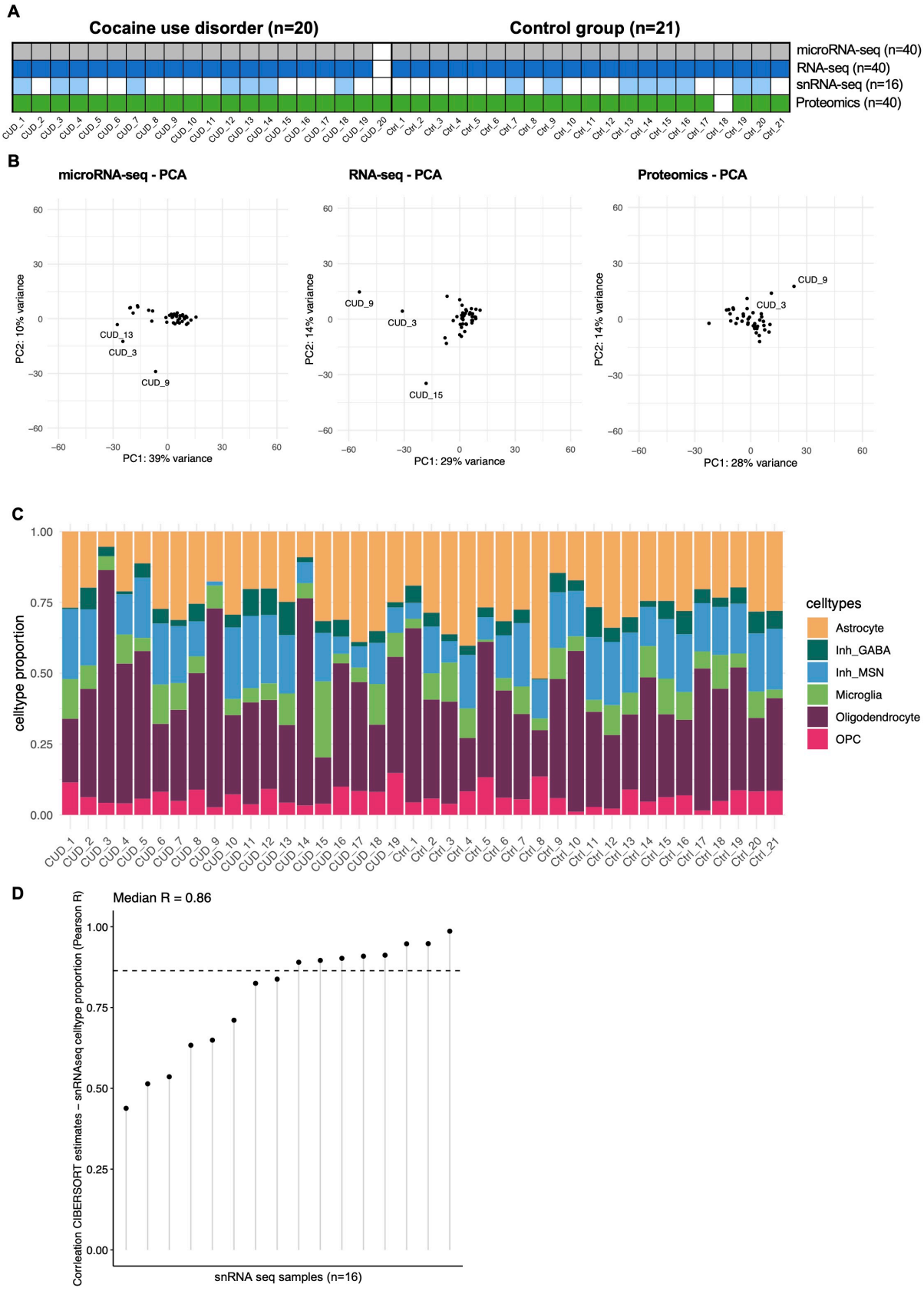


Figure S4.1 – Bulk dataset overview and cell type deconvolution analysis **A** comparative overview on dataset availability in the n=41 postmortem brain tissue samples of the ventral striatum (VS). **B** inspection and quality control of bulk datasets by principal component analysis in miRNA-seq, RNA-seq and proteomic datasets. **C** results of a cell type deconvolution analysis (CIBERSORT) for bulk RNA-seq data indicating estimated proportions of cell types in bulk RNA-seq samples. The reference dataset for the VS was generated using cell type specific

expression information from the n=16 snRNA-seq samples. Cell type deconvolution analysis was restricted to major cell types of the VS where DRD1- and DRD2-expressing medium spiny neurons (MSMs) were assigned to an Inh_MSN category and other non-MSN GABAergic interneurons were condensed to an Inh_GABA category. **D** correlation analysis between true cell type proportions identified by snRNA-seq and inferred proportions from CIBERSORT for the n=16 VS samples for which both bulk RNA-seq and snRNA-seq data was available.

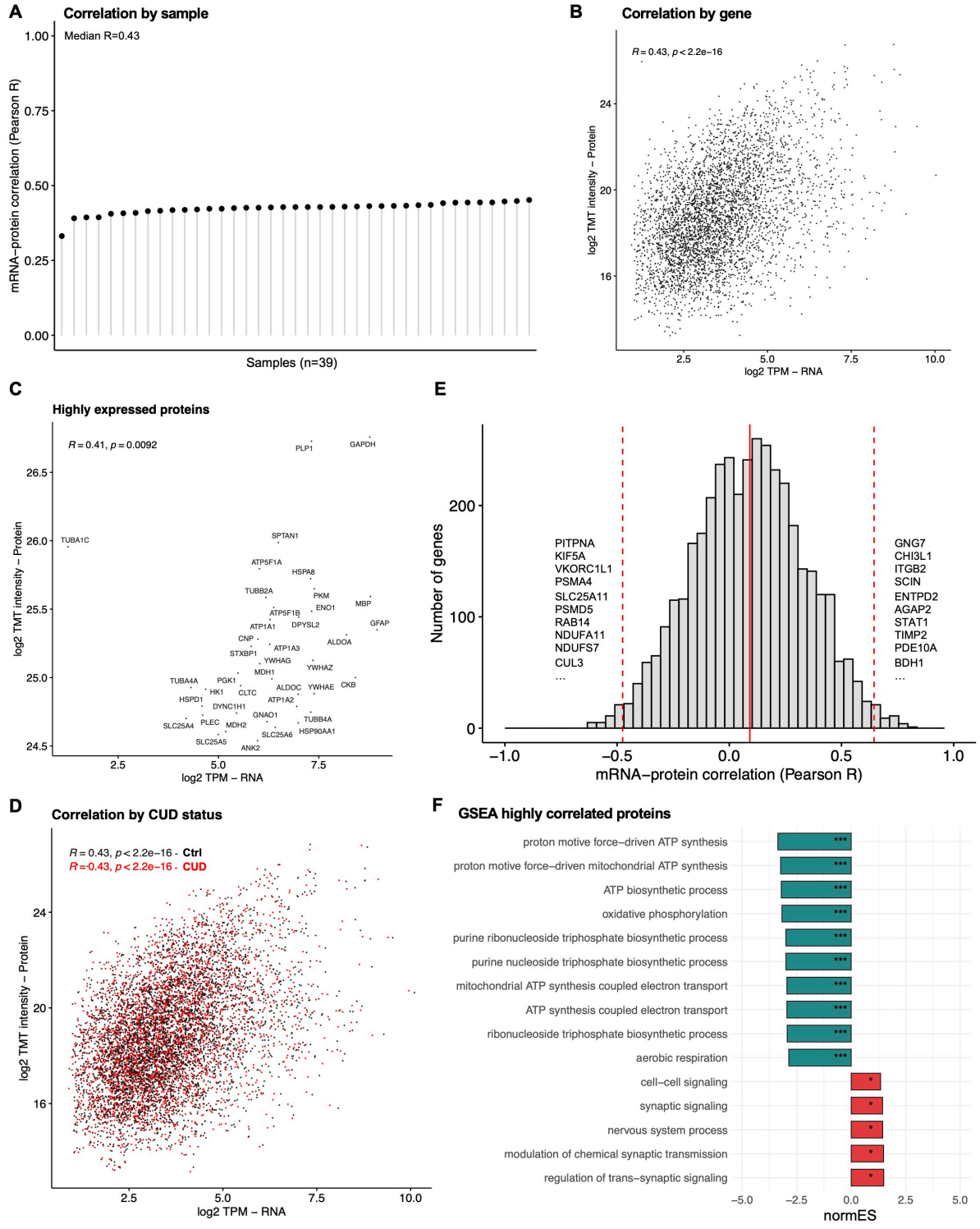


Figure S4.2 - Correlation analysis between transcriptome and proteome in the ventral striatum

A sample-level correlation analysis between transcriptome and proteome using n=3,935 genes for which RNA and protein expression levels were available. The analysis was performed in n=39 ventral striatum samples for which both transcriptomic and proteomic data was available. **B** correlation analysis of mean RNA (log2-transcript per million, log2 TPM) and protein expression values (log2 TMT-intensities) for each of the n=3,935 genes across samples. **C** correlation analysis of mean expression levels for RNA and proteins after selecting for highly expressed proteins (99th percentile of log2 TMT intensities). **D** sensitivity analysis of the correlation approach separated by CUD status. **E** histogram of Pearson correlation coefficients from a gene-centered approach

suggesting genes with strong positive (right) and negative (left) correlation of RNA and protein expression levels. RNA and protein expression levels were correlated individually for each gene without prior averaging of expression levels across samples (see STAR Methods). **F** pre-ranked gene-set enrichment analysis (GSEA) using gene-level correlation coefficients from **E** as the ranking metric. GSEA results separated by normalized enrichment scores (normES) suggest pathways with statistically significant enrichment for genes with strong positive (red) and negative correlation between RNA and protein expression levels. *p<0.05, **p<0.01, ***p<0.001.

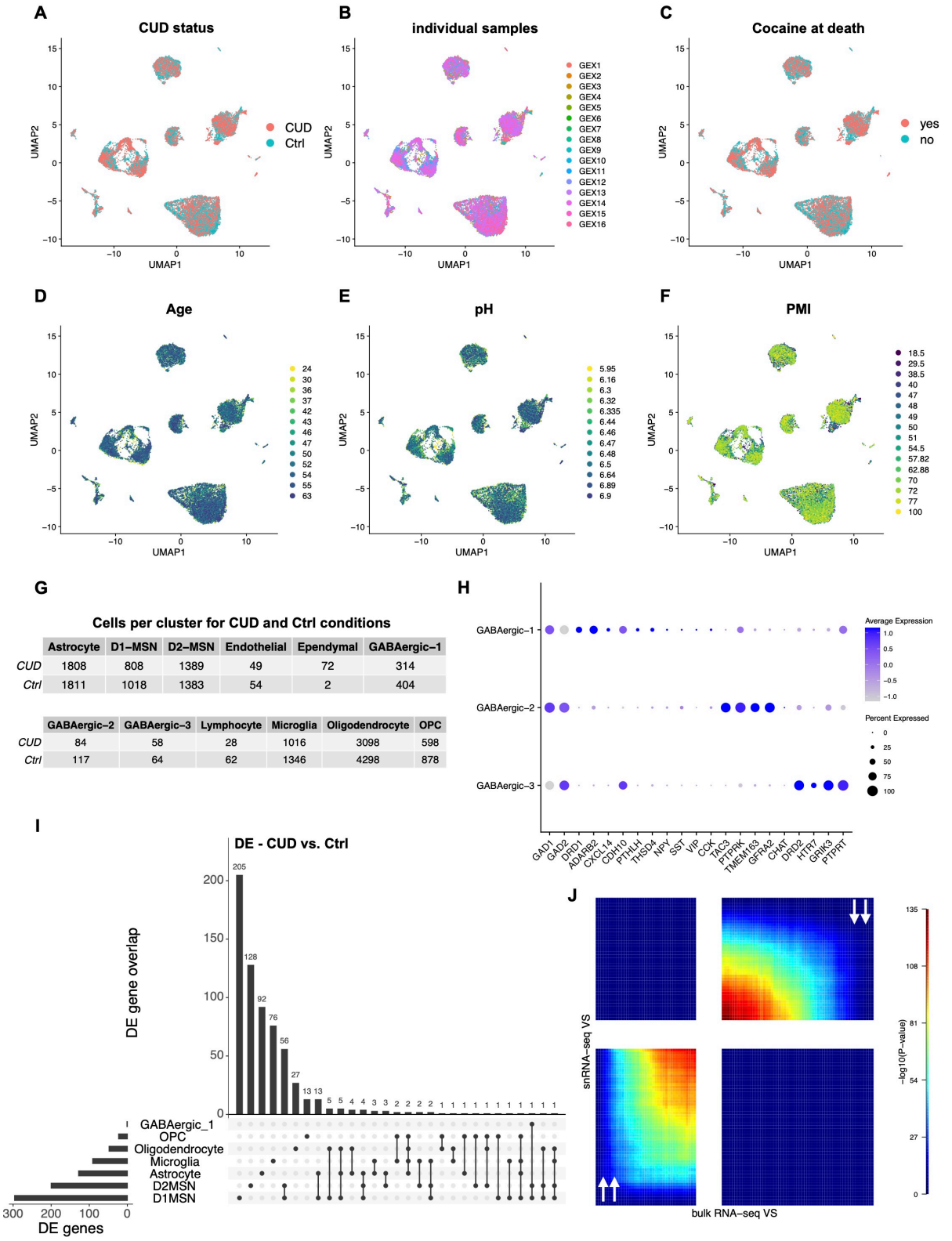


Figure S4.3 – Analysis of cell type-specific RNA expression profiles in n=20,759 single nuclei of the ventral striatum in cocaine use disorder
For the 12 distinct cell type clusters identified by single-nuclei (sn)RNA-seq in n=16 VS samples, the association of clustering patterns with phenotypes was assessed. **A** CUD status (CUD/Ctrl), **B** n=16 snRNA-seq samples (GEX1-GEX16), **C** cocaine at death status (yes/no), **D** tissue donor age (years), **E** brain tissue pH value, **F** postmortem interval (hours). **G** summary table depicting the number of retained nuclei per cell type in the snRNA-seq dataset for both CUD and Ctrl conditions. **H** to better characterize neuronal subpopulations in the ventral striatum, the expression of striatal

interneuron marker genes in clusters GABAergic-1, GABAergic-2, and GABAergic-3 was investigated. **I** the overlap between the n=653 differentially expressed genes (|log2FC| > 0.5, q<0.001) from 7 major cell types of the VS was visualized in an upset plot. **J** rank-rank hypergeometric overlap analysis of CUD-associated differential expression patterns from bulk RNA-seq in n=38 samples and snRNA-seq in n=16 samples of the ventral striatum (VS). Arrows in RRHO panels indicate convergent upregulation (bottom-left) or downregulation (top-right) of expression patterns across datasets.

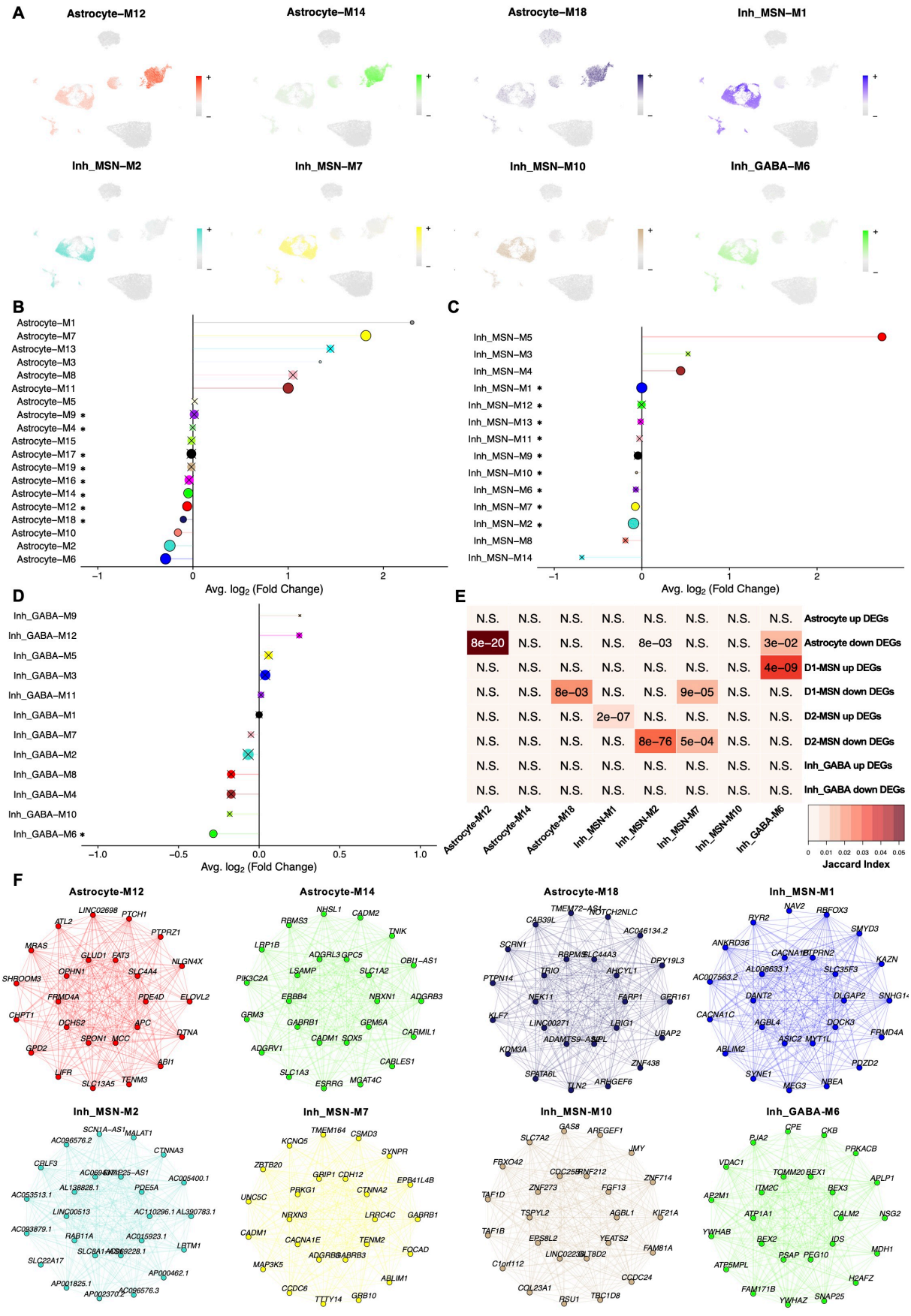


Figure S4.4 – Characterization of cell type specific co-expression networks identifies medium spiny neuron- and astrocyte-specific modules in cocaine use disorder Co-expression networks in the snRNA-seq dataset were identified for major cell types using

hdWGCNA. **A** module eigengene expression of cell type specific co-expression modules in CUD-associated astrocyte, Inh_GABA, and Inh_MSN clusters. Differential module eigengene expression analysis in CUD for **B** astrocytes, **C** medium spiny neurons (MSN, D1+D2 combined), and **D** non-MSN inhibitory GABAergic neurons (Inh_GABA). Co-expression modules were prioritized based on significant DME in the CUD vs. Ctrl comparison (colored dots with X) and cell type specificity (*) measured by module eigengene expression in the respective cell types. **E** overlap analysis between module genes from CUD- associated cell type specific co-expression modules and differentially expressed (DE) genes (|log2FC| > 0.5, q<0.001) from cluster-specific DE analysis. Jaccard index and p-values from Fisher-Test are shown. N.S.: not significant. **F** network plots of top 25 co-expression module hub genes for all cell type specific co-expression modules with significant DME in CUD.

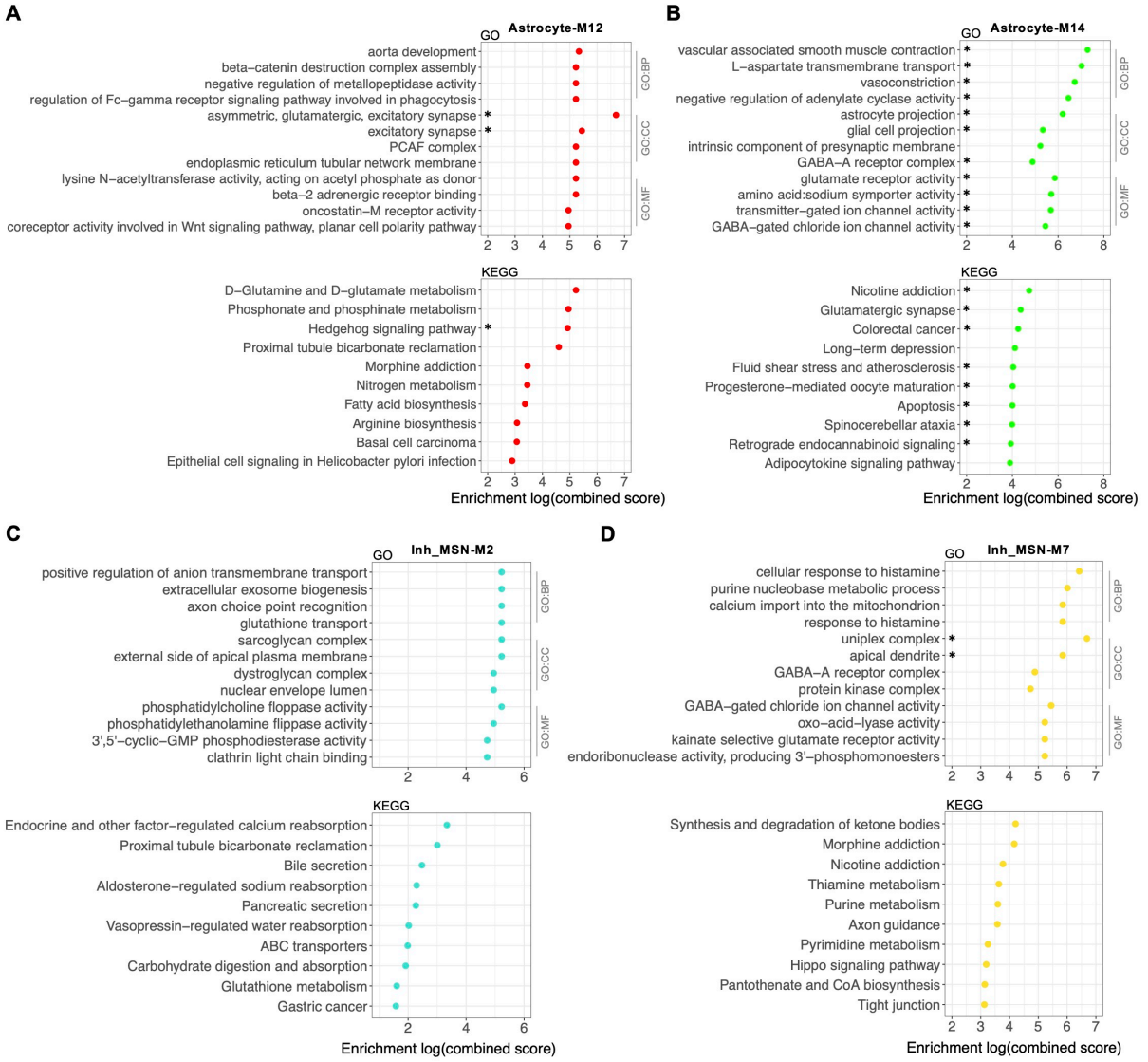


Figure S4.5 – Pathway enrichment analysis in astrocyte- and neuron-specific co-expression modules identifies altered glutamatergic and cAMP signaling in astrocytes and metabolic changes in medium spiny neurons
Results of GO and KEGG pathway enrichment analyses in co-expression modules **A** Astrocyte-M12, **B** Astrocyte-M14, **C** Inh_MSN-M2, and **D** Inh_MSN-M7. The combined score metrics for results ranking was derived from the enrichment statistics in EnrichR. For GO terms, the top 4 pathways from each ontology (GO Biological Process, GO:BP; GO Cellular Component, GO:CC; and GO Molecular Function, GO:MF) are shown. *FDR q<0.05.

Supplementary Tables

Due to genome-wide table sizes, for full supplementary tables, please refer to the online version of the supplementary information:

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/75306510-fc6c-4efc-a74d-5b194b00d965/mmc2.xlsx

Table S4.1. Demographic information for the full cohort of n = 41 tissue donors and the subset of individuals with n = 16 snRNA-seq data analyzed in the present study, related to Figures 4.1 and S4.1

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/f926580a-290e-4bae-a29b-f6e0cab5dc4e/mmc3.xlsx

Table S4.2. Bulk-level association analysis results of CUD in miRNA-seq, RNA-seq, and proteomics datasets, related to Figures 4.1 and S4.1. (A) Cell-type estimates in the bulk-level RNA-seq data using CIBERSORT for deconvolution analysis. DE summary statistics for CUD vs. Ctrl comparison in bulk-level analyses of (B) miRNA-seq, (C) RNA-seq, and (D) TMT-proteomics datasets, all related to Figure 1. (E) Statistically significant GO and KEGG pathway enrichment analysis results for RNA and protein associations with CUD. Enrichment analysis results were filtered for FDR q < 0.05.

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/e3aa07df-ffb8-442d-8150-47a3f461351b/mmc4.xlsx

Table S4.3. Transcriptome-proteome correlation analysis based on logTPM (RNA) and log2-TMT (proteomics) intensity values, related to STAR Methods and Figure S4.2. Left, sample-level correlation coefficients; middle, gene-level results of RNA/protein level correlation; right, correlation of RNA/protein level in dependence on CUD status. ΔR refers to the difference of correlation coefficients at the gene level between CUD and Ctrl.

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/20069a6b-8f07-449e-b245-d2fbcadca23a/mmc5.xlsx

Table S4.4. Results of bulk-level multi-omics analyses of CUD in the VS using WGCNA and MOFA, related to Figures 4.2 and 4.3. (A) Results from regression analysis for CUD in RNA and protein modules while adjusting for covariates and technical factors. Results are presented for the linear model module eigengene (ME) ~ CUD + Age + postmortem interval (PMI) + pH + RNA integrity number (RIN) (RNA) or batch (protein), ME column represents effect size from regression, and p_ME the associated p value. Significant p values for CUD are highlighted in bold. (B) Gene information on CUD-associated WGCNA modules from the analysis of transcriptomic and proteomic datasets. (C) statistically significant results (FDR q < 0.05) of the GO pathway enrichment analysis results for CUD-associated RNA and protein module genes from WGCNA. (D) Feature weights of miRNAs, RNAs, and proteins on CUD-associated factor 10 identified by MOFA. (E) Statistically significant GSEA results (FDR q < 0.05) of RNA and protein weights on CUD-associated MOFA factor 10.

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/62c6a9d8-4ca8-493f-b7bc-59d719153dc3/mmc6.xlsx

Table S4.5. Results of the cell-type-specific analysis of CUD-associated DE signatures using snRNA-seq, related to Figures 4.4 and S4.3. (A) Cell-type-specific differential gene expression results for major cell types of the VS (CUD vs. Ctrl comparison in n = 20,759 single nuclei) filtered

for |log2FC| > 0.5 and Bonferroni *p*_val_adj < 0.001. (B) Statistically significant GO enrichment analysis results (FDR q < 0.05) based on DEGs with |log2FC| > 0.5 and *p*.adj < 0.001.

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/c65842eb-8113-4e2e-8a56-5c744a63575c/mmc7.xlsx

Table S4.6. Cell-type-specific co-expression network analysis results from hdWGCNA, related to Figures 4.5, S4.4, and S4.5. (A) Identification of cell-type-specific co-expression modules in CUD by investigating DMEs and cell-type specificity of ME expression. (B) Gene information for cell-type-specific hdWGCNA modules with significant DME differences in the CUD vs. Ctrl comparison. (C) Statistically significant GO and KEGG enrichment analysis results (FDR q < 0.05) for cell-type-specific hdWGCNA modules associated with significant DME difference in CUD. (D–F) The analysis results in the same order from the investigation of the consensus modules between human CUD and a rodent model of chronic cocaine exposure.

https://www.cell.com/cms/10.1016/j.celrep.2025.115332/attachment/db43f867-c0a9-44b7-b5ce-5d273561fa0c/mmc8.xlsx

Table S4.7. Results of the astrocyte-neuron crosstalk analysis using CellChat, related to Figure 4.6. CUD vs. Ctrl comparisons were performed for LR pairs in astrocytes, D1-MSNs, and D2-MSNs to identify CUD-associated changes in the crosstalk between these major cell types of the ventral striatum.

5 DISCUSSION

The overarching aim of the three presented studies was the investigation and characterization of CUD-associated molecular signatures in postmortem human brain tissue. The first study examined genome-wide DNAm levels in more than 850,000 CpG sites in the human PFC and identified 20 DMRs that show a CUD-associated differential methylation pattern. For two of them, annotated to the genes *KLRN* and *NPFFR2*, a functional role of their gene product in modulating the behavioral response to cocaine was reported in rodent models, suggesting *KLRN* and *NPFFR2* as important genes in the pathophysiology of CUD. Network analyses identified CUD-associated co-methylation patterns among genes involved in neurotransmission and synaptic signaling, highlighting CUD-associated epigenetic deregulation of genes involved in neuronal functions. When performing integrative analyses of epigenomic and transcriptomic data in the PFC, as reported in the second study, these processes were confirmed to be also deregulated at the RNA level. At the same time, additional metabolic deregulation was identified, particularly of fatty acid metabolism in CUD. Two gene candidates were identified in CUD, *INPP5E* and *ZBTB4,* for which consistent associations with CUD were detected at all investigated omics levels, including genomic, epigenomic, transcriptomic, and transcriptome-wide splicing datasets. Using transcriptomic data from the PFC, the drug-repurposing analysis revealed glucocorticoid receptor-modulating drugs as a promising pharmacotherapeutic option that should be further investigated in CUD. Finally, miRNA and RNA transcriptomic, proteomic, and single-nuclei transcriptomic signatures of CUD were profiled in the ventral striatum, presented in the third study. Convergent evidence for the deregulation of metabolic pathways, particularly of fatty acid metabolism and oxidative phosphorylation, was identified in CUD, while glutamatergic signaling alterations were an additional important finding in the VS. By performing snRNA-seq, previously identified transcript deregulation patterns in CUD were assigned to individual cell types, revealing important cell type-specific molecular alterations of the VS in CUD.

5.1 Molecular characterization of cocaine use disorder in the human brain

5.1.1 Connecting the results between prefrontal cortex and ventral striatum

Strong connectivity between the PFC and striatal areas has been identified in the human brain, which is summarized under the concept of frontostriatal circuits (Morris et al., 2016). Of particular relevance in addiction are the glutamatergic projections from frontal areas to the striatum that are, for example, involved in executive control (Goldstein & Volkow, 2011). Thus, one can assume that functional alterations in the PFC can directly affect the function of striatal areas and vice versa, emphasizing the importance of a combined analysis of these two

brain regions in CUD. Multi-omics analyses in the PFC and the VS in a cohort of N=42 postmortem human brain donors were performed to characterize molecular alterations in these two important brain regions of the neurocircuitry of addiction. The three studies revealed a set of conserved results across regions that should be further discussed in detail.

5.1.2 Cocaine use disorder – a metabolic disorder of the brain?

One of the most prominent findings among transcriptional, but also protein-level associations with CUD were metabolic changes of the brain that were conserved across both the PFC and the VS. In particular, shared metabolic deregulation across brain regions was identified for genes involved in fatty acid metabolism and oxidative phosphorylation. In snRNA-seq of the VS, cell type-dependent directions of oxidative phosphorylation transcript deregulation were shown, and the same transcripts of oxidative phosphorylation genes were downregulated in neuronal but upregulated in glial cells. Therefore, one could assume that CUD, at least in part, resembles a metabolic disorder of the brain. Previous studies have indeed shown a pronounced effect of cocaine on cellular metabolism. For instance, cocaine was shown to significantly reduce brain glucose metabolism in multiple brain regions, according to a PET study in rodents, and this effect was shown to be dependent on cocaine's inhibitory effect on DAT (Thanos et al., 2008). Further, expression changes of genes involved in mitochondrial respiration were repeatedly observed in the brain at both the RNA and protein levels (Hámor et al., 2020; Zhou et al., 2011). A recent spatial multi-omics study investigated the acute effects of cocaine on different addiction-related brain regions in mice. Among the main findings of metabolomic analyses was the inhibition of the tricarboxylic acid cycle, while lipidomics indicated differential abundances of fatty acid species, suggesting a cocaine-induced effect on lipid homeostasis (Nezhyva et al., 2024). Experiments in cortical neuron cultures revealed a disrupted mitochondrial membrane potential and reduced ATP levels upon acute cocaine exposure (Cunha-Oliveira et al., 2006). In contrast, an inverse effect was observed in astrocytes with increased aerobic respiration and ATP production through a cocaine-induced metabolic switch from glycolysis to fatty acid oxidation (Natarajaseenivasan et al., 2018). These findings suggest important cell-type-specific effects of cocaine on oxidative phosphorylation. Mechanistically, reduced activity of respiratory chain enzymes, particularly of complex I, the NADH-dehydrogenase, was observed in response to cocaine exposure (Cunha-Oliveira et al., 2013; Devi & Chan, 1997). Overall, these cocaine effects result in increased levels of apoptosis, either by direct effects on pro-apoptotic pathways or by increasing the levels of reactive oxygen species, reducing ATP levels, and inducing autophagic cell death in the brain (Thornton et al., 2021).

As the analyses in BA9 and the VS confirmed transcriptional and proteomic changes in oxidative phosphorylation and fatty acid metabolism, and results were

in part also conserved in the CN, as shown by another recent study (Zillich et al., 2025), there is substantial evidence for an important role of metabolic deregulation in the pathophysiology of CUD. However, it must be noted that postmortem brain tissue only allows for cross-sectional analyses, making it difficult to disentangle the direct effects of cocaine from long-term neuroadaptations of metabolism as a consequence of chronic cocaine intake. While PMI was included as a covariate in statistical analysis to adjust for postmortem effects, degradation processes in postmortem brain tissue might themselves interfere with metabolism, thus affecting the results of association analyses with CUD. Thus, while there is substantial evidence for metabolic deregulation in the CUD brain, these findings have to be further studied, ideally on the longitudinal scale and in individual types.

5.1.3 The glutamate neurotransmitter system in cocaine use disorder

Across the methylation and expression analyses in BA9 as well as in the multi-omics analysis in the VS, glutamatergic signaling alterations emerged as a conserved finding in pathway analyses. Further, the cell-cell crosstalk analysis based on snRNA-seq data in the VS revealed altered glutamate signaling involving glutamate transporter *SLC1A2* as well as ionotropic and metabotropic glutamate receptors. The glutamate system is the main excitatory neurotransmitter system of the human brain. While cocaine's direct effect is related to dopamine, serotonin, and noradrenaline neurotransmission via reuptake inhibitory action on DAT, SERT, and NET, cocaine also indirectly increases extracellular glutamate levels, for instance, in the NAc (Smith et al., 1995). This effect was shown to be particularly expressed in animals that were repeatedly exposed to cocaine and have developed cocaine-related cue associations (McFarland et al., 2003). Here, overall lower baseline levels of extracellular and presynaptic glutamate are observed to rise strongly upon cocaine-induced glutamate release from prefrontal fibers terminating in the NAc (Schmidt & Pierce, 2010). Increased synaptic glutamate activates ionotropic receptors such as AMPA and NMDA receptors, as well as metabotropic mGluRs on the postsynaptic membrane. Activation of AMPA and NMDA receptors leads to increases in postsynaptic calcium concentrations and is involved in drug-induced neuroadaptations due to short-term and long-term neuroplastic changes (Lüscher & Malenka, 2011). Next to the relevance of glutamate neurotransmission in inducing neuroplastic changes during the intoxication stage, a strong relationship between glutamate levels in the human brain and craving was observed. For instance, cue-induced cocaine craving was associated with increased glutamate levels in the NAc, while a baseline reduction of glutamate levels was observed (Engeli et al., 2021), in line with results from rodent studies. These findings suggest modulators of the glutamate system as a potential pharmacotherapy for CUD. Pharmacological drugs targeting glutamate neurotransmission have indeed been tested in CUD, such as acamprosate, which reduces alcohol craving but did not have a significant effect on cocaine craving

(Kampman et al., 2011). In contrast, antagonizing mGluR5 showed promising results on reducing cocaine intake in a recent phase II trial (Gomez-Mancilla et al., 2025). Also, N-acetylcysteine, a prodrug that delivers cystine to the brain to promote the cystine-glutamate exchanger, showed promising results in preventing relapse in CUD (Nocito Echevarria et al., 2017). In summary, the glutamate system depicts a promising neurotransmitter system in CUD that should be further investigated to validate existing promising targets and identify new candidates that could be used in novel pharmacotherapies for CUD.

5.1.4 Cocaine use disorder and neuroinflammation

A conserved finding in bulk-level RNA-seq analyses of BA9 and VS were immunological signaling pathways that emerged in GO enrichment analyses on DEGs. These were related to leukocyte activation, migration, and cytokine signaling. Further, CUD-associated differential immune signaling was found in the snRNA-seq dataset focusing on individual cell populations. Here, specifically in microglia, a major cell type involved in immune processes in the brain, CUD-associated differential expression of genes involved in pathways related to T-cell receptor binding and MHC-related processes was observed. Thus, one can assume that CUD is associated with an altered neuroimmune response in the brain, which might be mediated by an altered microglia function. Further supporting the relevance of immune processes in CUD, drug repurposing analysis in BA9 identified glucocorticoid receptor-modulating drugs to revert the CUD-associated expression profile. The effects of cocaine on neuroinflammation have been intensely studied in recent years, but conflicting results have been reported. While in SUDs such as AUD or OUD, increased neuroinflammation in the brain was reported; a reduced neuroinflammatory state was hypothesized in CUD according to a bulk RNA-seq study that found downregulation of a large number of cytokines, including IL-6 and IL-17, in the CUD brain (Mews et al., 2023). However, the investigation of neuroinflammation in human postmortem brain tissue has significant limitations due to inherent contamination with immune cells from blood and cytokines, which often have context-dependent effects on inducing or reducing inflammatory responses. Since several studies have also suggested increased neuroinflammation in CUD (Correia et al., 2020), the picture is less clear, and based on cocaine's neurotoxic effects and its ability to increase levels of reactive oxygen species, cocaine-induced neuroinflammation is quite possible. This is supported by studies that found microglia activation upon repeated cocaine exposures in mice specific to the striatum with increases in IL-1β and TLR4, while no microglia activation was observed in cortical regions (Burkovetskaya et al., 2020). A direct effect of cocaine on TLR4 was reported, a toll-like receptor involved in the innate immune response that is strongly expressed in microglia (Northcutt et al., 2015). Activation of TLR4 signaling involves activation of NFκB, IL-1β, IL-6, and TNFα, indicating the induction of a pro-inflammatory state in the brain

(Brown et al., 2018; Russo et al., 2009). Inhibition of TLR4 signaling by the TLR4 antagonist (+)naltrexone impaired cocaine-conditioned place preference in a rat model, suggesting inflammatory signaling to modulate the reinforcing effect of cocaine (Northcutt et al., 2015). In summary, these results indicate that the role of neuroinflammation in CUD remains an important area of research, and future studies are needed to evaluate whether inflammation-modulating drugs, such as those targeting the glucocorticoid receptor, can be used as potential pharmacotherapy for CUD.

5.2 Methodological scopes and limitations

5.2.1 Composition of the discovery cohort

Postmortem human brain tissue originated from the Douglas Bell Canada Brain Bank (DBCBB), which comprises more than 3500 human brains and thereby provides essential resources for studies in neurological and psychiatric disorders. In the available set of brain tissue from donors with a diagnosis of CUD, pairwise matching based on age, sex, and the co-occurrence of other psychiatric disorders such as MDD and AUD was performed. Collections of postmortem brain tissue with SUD phenotypes are often biased towards male tissue donors due to sex-specific differences in the prevalence of the disorder and circumstances of death. While the inclusion of female tissue donors in the discovery cohort was intended, this was ultimately not possible, as the available sample size did not allow for sufficiently powered analyses in females. Also, a sufficient number of tissue donors could only be obtained from one ethnic background, namely the EA ethnicity. Thus, the discovery cohort can be considered relatively homogeneous as it consists of N=21 male CUD/Ctrl pairs, all of whom have an EA background.

During data analysis, a homogenous discovery cohort has the advantage of limiting the substantial inter-individual variation that is particularly evident in human samples due to different genomic but also environmental effects. While limited heterogeneity is beneficial for a systematic comparison between subjects that can differ in many variables, the sample of N=42 will most likely not represent the full spectrum of CUD. Thus, additional sex-specific analyses and investigations in different ethnic backgrounds are required as soon as sufficient sample sizes are available. Further, it has to be noted that the CUD phenotype is often not characterized by the excessive intake of cocaine only but also by a chronic intake of other drugs of abuse, such as alcohol, cannabis, and opioids (Stiltner et al., 2023). CUD-associated molecular signatures might thus reflect polysubstance use associations more, depending on the composition of the investigated sample. While this could be considered a critical issue in CUD cohorts with co-occurrent SUD diagnoses, examining a pure CUD phenotype characterized only by cocaine use could also be considered artificial, as this reflects only around 1% of CUD cases

(Stiltner et al., 2023). Subgroup analyses could be performed to address this issue in larger cohorts and meta-analyses.

5.2.2 Analyses in prefrontal cortex and ventral striatum

The selection of investigated omics for the studies in PFC and VS was strongly guided by their known biological relationships. For instance, DNAm is a known epigenetic regulation mechanism for gene expression (Gibney & Nolan, 2010), so RNA-seq was performed together with DNAm profiling in the PFC. Another regulatory relationship exists between miRNAs and mRNAs, as miRNAs regulate transcript abundance by a sequence-dependent induction of mRNA degradation. Parallel miRNA-seq and RNA-seq in the VS were performed to investigate their regulatory relationship. Further, mRNA is required for translation into proteins. Thus, mRNA expression is an important prerequisite for protein expression, which was the rationale behind the combined RNA-seq and proteomics analyses in the VS. In principle, all omics could have been profiled in the same brain regions, allowing unparalleled insights into the multi-omics signatures of CUD in the human brain. However, during quality control of the PFC samples, low RNA integrity numbers (RINs) were observed in a substantial fraction of the discovery sample, resulting in only N=25 of the N=42 samples that were available for RNA sequencing in BA9. Integrative multi-omics analyses would thus have suffered from a substantial number of missing data in the transcriptomic dataset, strongly reducing statistical power. Thus, a larger multi-omics analysis was performed in the VS, as the RIN value was consistently above critical thresholds in this brain region. The procedure for protein lysate production required a comparatively large amount of tissue, which is why a decision had to be made between the analysis of DNAm and protein levels in the VS. The analysis of the proteome was preferred, as proteins were so far underrepresented in previous studies on SUDs but depict the biological level that contributes significantly to brain function due to protein-protein interactions.

5.2.3 Comparability between brain regions and omics

An important prerequisite for informative comparative analyses of molecular signatures of SUDs in different brain regions is the availability of tissue samples derived from the same individuals. This is essential because the genetic background and environmental influences vary between individuals, and these factors are expected to have a substantial influence on disease associations and therefore also on comparative analyses. Tissue from BA9 and the VS originated from the same N=42 individuals, thus meeting the requirements. Further, cross-region analyses should rely on the same sequencing or analysis method to reduce methodology-associated biases. Direct comparability between BA9 and VS is possible with RNA-seq data, which is available at the bulk level in both brain regions. Comparison of results from GO enrichment analyses in DEGs and

WGCNA modules suggests deregulation of fatty acid metabolism, glia-related pathways, and immune processes as shared CUD-associated transcriptomic signatures in BA9 and the VS. One caveat in the direct comparison of RNA-seq data between BA9 and VS are potential batch effects, as the samples were not randomized prior to sequencing. To perform integrative analyses between regions, an in-parallel sample processing pipeline that also includes randomization would be required to ensure the elimination of batch effects introduced by sample handling.

Another comparison performed between methods was the analysis of transcriptomic signatures in the VS measured by bulk RNA-seq and snRNA-seq. A strong convergence of CUD-associated DEG patterns was observed by RRHO, suggesting that both methods are able to capture a similar set of disease associations. Still, snRNA-seq allows to assess the cell type specificity of findings. These results provide an important insight: integration approaches between bulk RNA-seq and snRNA-seq are, in principle, possible and can be used to assign already known bulk-level disease associations to individual cell types, as shown, for instance, for the oxidative phosphorylation changes in CUD.

As CUD was shown to be a substantially heritable disorder (Kendler et al., 2000; Kendler et al., 2007), integrating genomic data with other omics depicts an important milestone towards a comprehensive molecular characterization of CUD. While genetic information was used to characterize gene-level enrichment of CUD-associated genetic variants in gene candidates *INPP5E* and *ZBTB4* identified in BA9, no systematic integration analysis of SNP data with epigenomic, transcriptomic, and proteomic data was performed due to the small sample sizes and therefore limited statistical power of the CUD GWASs that are currently available. For a valid SNP association input, well-powered GWAS are required that will become available in the future (Deak et al., 2023).

5.2.4 Postmortem brain tissue

Postmortem brain tissue is currently the most important resource for investigating SUD-associated molecular alterations in the human brain. While, in principle, samples from the living brain could be available, for instance, when stereotactic surgery is performed, this is mainly restricted to disorders where surgeries are performed during treatment, for instance, during implantation of deep brain stimulation in Parkinson's disease. A general limitation of postmortem brain tissue is that it reflects the endpoint of the disorder, and the manner as well as circumstances of death might have effects on the molecular profile that are difficult to model in association analyses. Systematic comparisons between living and postmortem brain tissue are rare. Still, some first studies are available that compared RNA-seq data from living and postmortem brains, even at single-nuclei

resolution (Collado-Torres et al., 2023; Liharska et al., 2023; Vornholt et al., 2024). Using large-scale bulk RNA-seq data from human PFC, Liharska et al. (2023) found differential expression of more than 80% of genes between living and postmortem brains and therefore questioned the value of postmortem tissue for serving as a resource to study disease mechanisms of neuropsychiatric disorders. In contrast, a re-analysis of the dataset concluded that differential expression patterns from Liharska et al. (2023) are heavily confounded by RNA degradation, and adjustment for variables such as RIN and PMI strongly reduced the number of DEGs, thus refuting the raised doubts on the value of postmortem brain tissue (Collado-Torres et al., 2023). In the snRNA-seq analysis of the PFC by Vornholt et al. (2024), between 40 and 70% of genes were differentially expressed between living and postmortem brains across clusters; however, again, the differential expression analysis was performed without adjusting for RIN or PMI. Based on transcriptomic signatures from living brain samples, the authors developed an adjustment method that can be used to derive a covariate that could be included in future postmortem studies to correct for transcriptional differences between living and postmortem brains.

5.3 Future perspectives

5.3.1 Advancing multi-omics

Traditionally, there has been a distinction between “genome-first” and “phenotype-first” approaches in multi-omics research (Hasin et al., 2017). In the genome-first approach, multi-omics is used to identify the functional relevance of causal GWAS variants, for example, by relating them to transcript or protein levels using quantitative trait loci analyses. Integration of information across omics levels could then allow the assignment of specific gene-, pathway-, and systems-level alterations to genetic variants that have a causative role in a disease. In contrast, the phenotype-first approach centers around the disorder phenotype without focusing on a specific gene locus. It tries to identify molecular signatures associated with this disorder to better understand its underlying disease processes. The three studies presented above can be assigned to the phenotype-first category. The integration of genetic information with other omics still depicts an important milestone in multi-omics analysis, as disorders often have a genetic but also an environmental component that both affect the epigenome, transcriptome, and proteome. Distinguishing molecular signatures of genetic vulnerability from those of environmental effects, as well as signatures that reflect disorder progression, remains a major methodological challenge. While in principle, this could be addressed by longitudinal multi-omics analyses, this is often challenging or even impossible in humans, depending on the tissue type. The development of complex modeling techniques such as MEFISTO (Velten et al., 2022), which enable longitudinal multi-omics analyses in addition to the integration of longitudinal

data from surrogate tissues such as blood or from rodent models, could represent approaches to overcome this problem.

One of the key concepts of multi-omics is the assumption that adding information from additional molecular levels allows a better and more comprehensive characterization of biological systems. Following this logic, the integration of the previously profiled omics, including genomics, epigenomics, transcriptomics, and proteomics, with additional omics, such as metabolomics, lipidomics, or glycomics depicts an important approach to gaining deeper insights into the molecular alterations associated with CUD. Here, innovative multi-omics methods and standardized analysis pipelines are required to carry out meaningful integration analyses with increasingly complex datasets.

5.3.2 Data synthesis approaches

While multi-omics studies themselves depict data synthesis approaches, there is a clear benefit of combining results from multi-omics analyses, for instance, in meta-analyses or in integrative approaches combining bulk-level with single-cell datasets. Meta-analyses are essential in identifying robust associations with CUD by substantially increasing the statistical power and enabling subgroup analyses stratified by sex, ethnicity, or comorbidities. As shown in the multi-omics study of CUD in the VS, integrating snRNA-seq with RNA-seq data is a feasible approach that provides novel insights into the cell type specificity of bulk-level associations. As single-cell analyses are still comparatively expensive, synthesis approaches that follow up robust findings from large-scale bulk-level meta-analyses in single-cell datasets could provide important insights into the role of individual cell types in disease mechanisms. Another valuable resource in addiction research are rodent models that reflect addiction-like criteria, such as the 3-crit model of cocaine addiction (Deroche-Gamonet et al., 2004). By functionally validating findings derived from human multi-omics analyses, such models are extremely valuable in selecting and prioritizing molecular targets, as addiction criteria such as persistence of cocaine-seeking, motivation for cocaine-taking, and resistance to punishment can be tested upon molecular genetic or pharmacological modulation of potential drug targets (Pohořalá et al., 2021). In addition, organoid models derived from human induced pluripotent stem cells represent another model system that could contribute to functional validation of findings, thereby contributing towards translation into the clinic (Davis et al., 2025).

5.4 Conclusion

Treatment options for CUD are currently limited, and even the most effective therapies report a large fraction of patients who relapse or do not remain adherent to therapy. In the multi-omics analyses focusing on two important brain regions in addiction, the PFC and the VS, molecular signatures of CUD were identified in the

postmortem human brain by combining information from genomic, epigenomic, transcriptomic, and proteomic datasets. By performing multi-omics analyses involving bulk and single-cell datasets together with replication analyses in independent datasets reflecting human CUD but also rodent models of cocaine addiction, converging evidence was identified for deregulation of oxidative phosphorylation, fatty acid metabolism, and glutamatergic signaling in the CUD brain. Following these results, meta-analyses, functional validation experiments, and advanced statistical modeling methods for multi-omics datasets will help expand the knowledge of the pathophysiology of CUD and pave the way for the development of novel therapeutic options.

6 SUMMARY

6.1 English Summary

Cocaine use disorder is a severe psychiatric disorder characterized by excessive and compulsive use of cocaine, as well as difficulties in reducing cocaine intake and remaining abstinent. Due to the chronic effects of cocaine on the organism, individuals with cocaine use disorder have a significantly increased risk for the development of cardiovascular disease and stroke, as well as comorbid psychiatric disorders such as depression, resulting in a substantial burden of disease. Several million people worldwide suffer from cocaine use disorder. Still, the currently available treatment approaches have limited effectiveness and are associated with high relapse rates, highlighting that the successful treatment of cocaine use disorder remains a major challenge. Regarding the etiology of the disease, it is assumed that changes in brain neurocircuits contribute to the development and maintenance of cocaine use disorder. In this context, the "neurocircuitry of addiction" model has been developed, which allows to relate the symptoms of cocaine use disorder to specific regions of the brain, such as the striatum or cortical areas. Molecular changes at the level of epigenetics, transcription, and protein expression are hypothesized to underlie the neuroadaptations and neurocircuit alterations in cocaine use disorder. Despite intensive research on the molecular underpinnings of cocaine use disorder, many aspects remain unclear, for instance, how changes in epigenetics are related to transcriptional and protein-level changes, while also the role of different cell types in the brain in cocaine use disorder is largely unknown. To characterize molecular alterations in the human brain, postmortem brain tissue serves as an important resource for investigating molecular changes using case-control study designs. The aim of the three studies presented was to conduct a comprehensive molecular characterization of cocaine use disorder in the human brain and to provide deeper insights into its pathophysiology, as well as potential new therapeutic approaches for cocaine use disorder.

In the first study, tissue from Brodmann Area 9, a subregion of the prefrontal cortex, from 21 individuals with cocaine use disorder and 21 control subjects was analyzed for genome-wide changes in DNA methylation. The second study represents a multi-omics approach performing integrative analyses of DNA methylation and gene expression data obtained by RNA sequencing. In addition, RNA splicing patterns in Brodmann Area 9 were examined, and the conservation of gene expression signatures in the prefrontal cortex in cocaine use disorder was investigated using replication datasets. The third study represents a comprehensive molecular characterization of cocaine use disorder in the same cohort, but in a striatal brain region, the ventral striatum. Omics-wide datasets of

microRNA, RNA, and protein expression were integrated, and cell type-specific transcriptomic changes were identified using single-nuclei RNA sequencing. To investigate the biological role of molecular changes associated with cocaine use disorder, pathway and network analyses were performed, as well as drug repurposing analyses to identify drugs that can reverse the molecular brain changes associated with cocaine use disorder. The epigenome-wide association study in Brodmann Area 9 identified a total of 20 differentially methylated regions associated with cocaine use disorder. In addition, network analyses revealed an enrichment of differential DNA methylation within genes involved in neurotransmission and synaptic signaling. Data integration with information on RNA expression, as performed in the second study, confirmed differential expression of synaptic genes at the RNA level and also suggested altered oxidative phosphorylation and fatty acid metabolism in the brain. With *INPP5E* and *ZBTB4*, two gene candidates were identified that showed significant changes in individuals with cocaine use disorder across analyses of DNA methylation, splicing, and RNA expression while also containing genetic risk variants. Pharmacological compounds that act on the glucocorticoid receptor were identified as potential drugs that could reverse the cocaine use disorder-associated gene expression profile. The study in the ventral striatum confirmed findings of altered oxidative phosphorylation and fatty acid metabolism in the brain of individuals with cocaine use disorder and represents one of the first studies that included proteomic profiling in multi-omics approaches for addiction research. The transcriptome analysis at the single nuclei level highlighted a particular relevance of striatal astrocytes and medium spiny neurons in cocaine use disorder, suggesting alterations in cell-cell adhesion and glutamatergic neurotransmission.

In summary, the three studies represent an in-depth molecular characterization of cocaine use disorder in postmortem human brain tissue, providing deep insights into the molecular changes in two important regions of the neurocircuitry of addiction: the prefrontal cortex and the ventral striatum. The integration of multiple omics-wide data sets, both at the bulk level and at the single nuclei level using complex biostatistical methods, highlights the additional value that results from integrative analyses of high-dimensional data sets, for example, allowing the identification of changes in RNA expression that are directly linked to altered DNA methylation. Promising targets for future therapeutic approaches in cocaine use disorder could, for example, be related to glutamatergic neurotransmission, fatty acid and mitochondrial metabolism, as well as at (neuro-)immunological targets such as the glucocorticoid receptor. Functional validation of findings from multi-omics studies, for example, using animal or organoid models, depicts the next step to use this knowledge for a better understanding of the underlying disease mechanisms and for developing novel therapeutic approaches for cocaine use disorder.

6.2 German Summary – Zusammenfassung in deutscher Sprache

Die Kokainkonsumstörung ist eine schwere psychiatrische Erkrankung, die durch einen übermäßigen und zwanghaften Konsum von Kokain, sowie durch Schwierigkeiten, den Konsum von Kokain zu reduzieren oder zu beenden, charakterisiert ist. Durch die chronische Einwirkung von Kokain auf den Organismus haben Personen mit Kokainkonsumstörung ein signifikant erhöhtes Risiko für die Entwicklung von Herz-Kreislauf-Erkrankungen und Schlaganfällen, sowie für weitere psychiatrische Erkrankungen wie Depression, was zu einer hohen Krankheitslast führt. Weltweit leiden mehrere Millionen Menschen an einer Kokainkonsumstörung. Momentan verfügbare Therapieansätze haben nur eine eingeschränkte Wirksamkeit und sind mit einer hohen Rate an Rückfällen verbunden. Eine erfolgreiche Behandlung der Kokainkonsumstörung stellt somit eine große Herausforderung dar. Aus ätiologischer Sicht tragen neuronale Veränderungen im Gehirn maßgeblich zur Entstehung und Aufrechterhaltung der Kokainkonsumstörung bei. In diesem Zusammenhang wurde das Modell der „neurocircuitry of addiction“ (neuronale Schaltkreise der Sucht) entwickelt, auf Basis derer die Symptome der Kokainkonsumstörung spezifischen Regionen des Gehirns, wie dem Striatum oder kortikalen Bereichen, zugeordnet werden können. Molekulare Veränderungen auf Ebene der Epigenetik, Transkription und Proteinexpression werden als Grundlagen der Neuroadaptation und der Veränderung von neuronalen Schaltkreisen bei der Kokainkonsumstörung im Gehirn angenommen. Trotz intensiver Forschung an diesen molekularen Grundlagen der Kokainkonsumstörung sind viele Aspekte noch größtenteils unklar, beispielsweise wie Veränderungen der Epigenetik mit Veränderungen auf RNA- und Proteinebene zusammenhängen, sowie welche Rolle unterschiedliche Zelltypen im Gehirn bei der Kokainkonsumstörung haben. Postmortales Gehirngewebe ist eine zentrale Ressource zur Charakterisierung molekularer Mechanismen im menschlichen Gehirn. Ziel der drei vorgelegten Studien war es, eine umfassende molekulare Charakterisierung der Kokainkonsumstörung im menschlichen Gehirn durchzuführen, deren Erkenntnisse umfassende Einblicke in die Pathophysiologie sowie in potentiell neue therapeutische Ansätze bei der Kokainkonsumstörung ermöglichen sollen.

In der ersten Studie wurde Gewebe aus dem Brodmann-Areal 9, einer Subregion des präfrontalen Cortex, von 21 Personen mit Kokainkonsumstörung und 21 Kontrollpersonen, auf genomweite Veränderungen der DNA-Methylierung hin untersucht. Die zweite Studie stellt eine integrative Multi-Omics Analyse dar, bei der die Daten der DNA-Methylierung mit durch RNA-Sequenzierung gewonnener Information zur Genexpression kombiniert wurden. Hierbei wurden zudem Splicing-Muster der RNA im Brodmann-Areal 9 untersucht sowie mittels Replikationsdatensätzen konservierte Genexpressions-Signaturen im präfrontalen Cortex bei Kokainkonsumstörung bestimmt. Die dritte Studie stellt

eine umfassende molekulare Charakterisierung der Kokainkonsumstörung in der gleichen Stichprobe, allerdings in einer striatalen Gehirnregion, dem ventralen Striatum, dar. Hierzu wurden omics-weite Datensätze der Expression von microRNA, RNA, und Proteinen integriert sowie mittels Einzelkern-RNA-Sequenzierung zelltypspezifische Unterschiede des Transkriptoms herausgearbeitet. Um die biologische Relevanz der kokainkonsumstörungs-assoziierten molekularen Veränderungen zu bestimmen, wurden Pathway- und Netzwerkanalysen durchgeführt sowie Analysen des Repurposings von Arzneimitteln, um Medikamente zu identifizieren, die den molekularen Veränderungen bei der Kokainkonsumstörung im Gehirn entgegenwirken können. Die epigenomweite Assoziationsstudie im Brodmann-Areal 9 identifizierte insgesamt 20 differentiell methylierte Regionen, die mit einer Kokainkonsumstörung assoziiert waren. Zudem zeigte sich in Netzwerkanalysen eine Anreicherung von differentieller DNA-Methylierung innerhalb von Genen, die in Neurotransmission und synaptische Aktivität involviert sind. Die Datenintegration mit Informationen zur RNA-Expression im Rahmen der zweiten Studie bestätigte die differentielle Expression von synaptischen Genen auch auf RNA-Ebene und wies zudem auf Veränderungen in der oxidativen Phosphorylierung und des Fettsäurestoffwechsels im Gehirn hin. Mit *INPP5E* und *ZBTB4* konnten zwei Gen-Kandidaten identifiziert werden, die über die Analysen der DNA-Methylierung, des Splicings und der RNA-Expression hinweg signifikante Unterschiede bei Personen mit Kokainkonsumstörung zeigten und zudem genetische Risikovarianten enthielten. Als potentielle Medikamente, die das mit der Kokainkonsumstörung assoziierte Genexpressionsprofil umkehren könnten, wurden Pharmaka identifiziert, die auf den Glukokortikoid-Rezeptor einwirken. Die Studie im ventralen Striatum bestätigte die Befunde der veränderten oxidativen Phosphorylierung und des Fettsäurestoffwechsels im Gehirn von Personen mit Kokainkonsumstörung und stellt eine der ersten Studien dar, die Veränderungen des Proteoms im Rahmen von Multi-Omics Ansätzen bei Suchterkrankungen mit einbeziehen. Die Analyse des Transkriptoms auf Einzelzellebene wies auf eine besondere Relevanz von Astrozyten und mittelgroßer dorntragender Projektionsneurone des Striatums bei der Kokainkonsumstörung hin, die auf Veränderungen im Bereich der Zell-Zell-Adhäsion und glutamaterger Neurotransmission hinweist.

Zusammenfassend stellen die drei Studien eine umfassende molekulare Charakterisierung der Kokainkonsumstörung im postmortalen menschlichen Gehirn dar, die tiefe Einblicke in die molekularen Veränderungen in zwei wichtigen Regionen der „neurocircuitry of addiction“ ermöglicht: im präfrontalen Cortex und im ventralen Striatum. Die Integration mehrerer omics-weiter Datensätze, sowohl auf Gesamtebene als auch auf Einzelzellebene, mittels komplexer biostatistischer Methoden zeigt den Mehrwert auf, der sich durch integrative Analysen hochdimensionaler Datensätze ergibt: Beispielsweise

können Veränderungen der RNA-Expression direkt mit denen der DNA-Methylierung in Verbindung gesetzt werden. Vielversprechende Ansätze für zukünftige Therapieansätze der Kokainkonsumstörung könnten beispielsweise in der glutamatergen Neurotransmission, im Fettsäure- und mitochondrialen Stoffwechsel, sowie an (neuro-)immunologischen Zielen wie dem Glukokortikoid-Rezeptor verortet sein. Die funktionelle Validierung der Befunde aus Multi-Omics-Studien, beispielsweise mittels Tier- oder Organoidmodellen, stellt den nächsten Schritt dar, um dieses Wissen für ein besseres Verständnis der zugrundeliegenden Krankheitsmechanismen und für die Entwicklung neuer Therapieansätze bei der Kokainkonsumstörung zu nutzen.

7 REFERENCES

Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, *17*(8), 487-500. https://doi.org/10.1038/nrg.2016.59

Ambrosi, C., Manzo, M., & Baubec, T. (2017). Dynamics and Context-Dependent Roles of DNA Methylation. *Journal of Molecular Biology*, *429*(10), 1459-1475. https://doi.org/https://doi.org/10.1016/j.jmb.2017.02.008

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Amrute, J. M., Luo, X., Penna, V., Yang, S., Yamawaki, T., Hayat, S., Bredemeyer, A., Jung, I. H., Kadyrov, F. F., Heo, G. S., Venkatesan, R., Shi, S. Y., Parvathaneni, A., Koenig, A. L., Kuppe, C., Baker, C., Luehmann, H., Jones, C., Kopecky, B., . . . Lavine, K. J. (2024). Targeting immune-fibroblast cell communication in heart failure. *Nature*, *635*(8038), 423-433. https://doi.org/10.1038/s41586-024-08008-5

Anier, K., Malinovskaja, K., Aonurm-Helm, A., Zharkovsky, A., & Kalda, A. (2010). DNA methylation regulates cocaine-induced behavioral sensitization in mice. *Neuropsychopharmacology*, *35*(12), 2450-2461. https://doi.org/10.1038/npp.2010.128

Anier, K., Zharkovsky, A., & Kalda, A. (2013). S-adenosylmethionine modifies cocaine-induced DNA methylation and increases locomotor sensitization in mice. *Int J Neuropsychopharmacol*, *16*(9), 2053-2066. https://doi.org/10.1017/s1461145713000394

Argelaguet, R., Arnol, D., Bredikhin, D., Deloro, Y., Velten, B., Marioni, J. C., & Stegle, O. (2020). MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biology*, *21*(1), 111. https://doi.org/10.1186/s13059-020-02015-1

Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J. C., Buettner, F., Huber, W., & Stegle, O. (2018). Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology*, *14*(6), e8124. https://doi.org/https://doi.org/10.15252/msb.20178124

Athieniti, E., & Spyrou, G. M. (2023). A guide to multi-omics data collection and integration for translational medicine. *Comput Struct Biotechnol J*, *21*, 134-149. https://doi.org/10.1016/j.csbj.2022.11.050

Bååth, R. (2014). Bayesian First Aid: A Package that Implements Bayesian Alternatives to the Classical *.test Functions in R. In the proceedings of UseR! 2014 - the International R User Conference.

Baker-Andresen, D., Zhao, Q., Li, X., Jupp, B., Chesworth, R., Lawrence, A. J., & Bredy, T. (2015). Persistent variations in neuronal DNA methylation following cocaine self-administration and protracted abstinence in mice. *Neuroepigenetics*, *4*, 1-11. https://doi.org/10.1016/j.nepig.2015.10.001

Baldaçara, L., Cogo-Moreira, H., Parreira, B. L., Diniz, T. A., Milhomem, J. J., Fernandes, C. C., & Lacerda, A. L. (2016). Efficacy of topiramate in the treatment of crack cocaine dependence: a double-blind, randomized, placebo-controlled trial. *J Clin Psychiatry*, *77*(3), 398-406. https://doi.org/10.4088/JCP.14m09377

Bannon, M. J., Johnson, M. M., Michelhaugh, S. K., Hartley, Z. J., Halter, S. D., David, J. A., Kapatos, G., & Schmidt, C. J. (2014). A Molecular Profile of Cocaine Abuse Includes the Differential Expression of Genes that Regulate Transcription, Chromatin, and Dopamine Cell Phenotype. *Neuropsychopharmacology*, *39*(9), 2191-2199. https://doi.org/10.1038/npp.2014.70

Baysoy, A., Bai, Z., Satija, R., & Fan, R. (2023). The technological landscape and applications of single-cell multi-omics. *Nature Reviews Molecular Cell Biology*, *24*(10), 695-713. https://doi.org/10.1038/s41580-023-00615-w

Belin, D., & Everitt, B. J. (2008). Cocaine Seeking Habits Depend upon Dopamine-Dependent Serial Connectivity Linking the Ventral with the Dorsal Striatum. *Neuron*, *57*(3), 432-441. https://doi.org/https://doi.org/10.1016/j.neuron.2007.12.019

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289-300. http://www.jstor.org/stable/2346101

Bernstein, E., Judith, B., Katherine, T., Anne, V., Timothy, H., Suzette, L., & and Hingson, R. (2006). Racial and Ethnic Diversity Among a Heroin and Cocaine Using Population. *Journal of Addictive Diseases*, *24*(4), 43-63. https://doi.org/10.1300/J069v24n04_04

Bierut, L. J., Strickland, J. R., Thompson, J. R., Afful, S. E., & Cottler, L. B. (2008). Drug use and dependence in cocaine dependent subjects, community-based individuals, and their siblings. *Drug and Alcohol Dependence*, *95*(1), 14-22. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2007.11.023

Biondich, A. S., & Joslin, J. D. (2016). Coca: The History and Medical Significance of an Ancient Andean Tradition. *Emerg Med Int*, *2016*, 4048764. https://doi.org/10.1155/2016/4048764

Birdsey, G. M., Shah, A. V., Dufton, N., Reynolds, L. E., Osuna Almagro, L., Yang, Y., Aspalter, I. M., Khan, S. T., Mason, J. C., Dejana, E., Göttgens, B., Hodivala-Dilke, K., Gerhardt, H., Adams, R. H., & Randi, A. M. (2015). The endothelial transcription factor ERG promotes vascular stability and

growth through Wnt/ß-catenin signaling. *Dev Cell*, *32*(1), 82-96. https://doi.org/10.1016/j.devcel.2014.11.016

Bogdan, R., Hatoum, A. S., Johnson, E. C., & Agrawal, A. (2023). The Genetically Informed Neurobiology of Addiction (GINA) model. *Nature Reviews Neuroscience*, *24*(1), 40-57. https://doi.org/10.1038/s41583-022-00656-8

Bonev, B., Castelo-Branco, G., Chen, F., Codeluppi, S., Corces, M. R., Fan, J., Heiman, M., Harris, K., Inoue, F., Kellis, M., Levine, A., Lotfollahi, M., Luo, C., Maynard, K. R., Nitzan, M., Ramani, V., Satijia, R., Schirmer, L., Shen, Y., . . . Nowakowski, T. J. (2024). Opportunities and challenges of single-cell and spatially resolved genomics methods for neuroscience discovery. *Nature Neuroscience*, *27*(12), 2292-2309. https://doi.org/10.1038/s41593-024-01806-0

Brenner, E., Tiwari, G. R., Kapoor, M., Liu, Y., Brock, A., & Mayfield, R. D. (2020). Single cell transcriptome profiling of the human alcohol-dependent brain. *Hum Mol Genet*, *29*(7), 1144-1153. https://doi.org/10.1093/hmg/ddaa038

Brown, K. T., Levis, S. C., O'Neill, C. E., Northcutt, A. L., Fabisiak, T. J., Watkins, L. R., & Bachtell, R. K. (2018). Innate immune signaling in the ventral tegmental area contributes to drug-primed reinstatement of cocaine seeking. *Brain, Behavior, and Immunity*, *67*, 130-138. https://doi.org/https://doi.org/10.1016/j.bbi.2017.08.012

Burkovetskaya, M. E., Small, R., Guo, L., Buch, S., & Guo, M. L. (2020). Cocaine self-administration differentially activates microglia in the mouse brain. *Neurosci Lett*, *728*, 134951. https://doi.org/10.1016/j.neulet.2020.134951

Butler, A. J., Rehm, J., & Fischer, B. (2017). Health outcomes associated with crack-cocaine use: Systematic review and meta-analyses. *Drug and Alcohol Dependence*, *180*, 401-416. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2017.08.036

Cabana-Domínguez, J., Shivalikanjli, A., Fernàndez-Castillo, N., & Cormand, B. (2019). Genome-wide association meta-analysis of cocaine dependence: Shared genetics with comorbid conditions. *Prog Neuropsychopharmacol Biol Psychiatry*, *94*, 109667. https://doi.org/10.1016/j.pnpbp.2019.109667

Cabrera-Mendoza, B., Stertz, L., Najera, K., Selvaraj, S., Teixeira, A. L., Meyer, T. D., Fries, G. R., & Walss-Bass, C. (2023). Within subject cross-tissue analyzes of epigenetic clocks in substance use disorder postmortem brain and blood. *Am J Med Genet B Neuropsychiatr Genet*, *192*(1-2), 13-27. https://doi.org/10.1002/ajmg.b.32920

Caffino, L., Messa, G., & Fumagalli, F. (2018). A single cocaine administration alters dendritic spine morphology and impairs glutamate receptor synaptic retention in the medial prefrontal cortex of adolescent rats. *Neuropharmacology*, *140*, 209-216. https://doi.org/https://doi.org/10.1016/j.neuropharm.2018.08.006

Cahill, K. M., Huo, Z., Tseng, G. C., Logan, R. W., & Seney, M. L. (2018). Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach. *Sci Rep*, *8*(1), 9588. https://doi.org/10.1038/s41598-018-27903-2

Camilo, C., Maschietto, M., Vieira, H. C., Tahira, A. C., Gouveia, G. R., Feio Dos Santos, A. C., Negrão, A. B., Ribeiro, M., Laranjeira, R., Vallada, H., & Brentani, H. (2019). Genome-wide DNA methylation profile in the peripheral blood of cocaine and crack dependents. *Braz J Psychiatry*, *41*(6), 485-493. https://doi.org/10.1590/1516-4446-2018-0092

Campbell, R. R., Chen, S., Beardwood, J. H., López, A. J., Pham, L. V., Keiser, A. M., Childs, J. E., Matheos, D. P., Swarup, V., Baldi, P., & Wood, M. A. (2021). Cocaine induces paradigm-specific changes to the transcriptome within the ventral tegmental area. *Neuropsychopharmacology*, *46*(10), 1768-1779. https://doi.org/10.1038/s41386-021-01031-4

Cannella, N., Oliveira, A. M. M., Hemstedt, T., Lissek, T., Buechler, E., Bading, H., & Spanagel, R. (2018). Dnmt3a2 in the Nucleus Accumbens Shell Is Required for Reinstatement of Cocaine Seeking. *J Neurosci*, *38*(34), 7516-7528. https://doi.org/10.1523/jneurosci.0600-18.2018

Ceceli, A. O., Huang, Y., Kronberg, G., Malaker, P., Miller, P., King, S. G., Gaudreault, P. O., McClain, N., Gabay, L., Vasa, D., Newcorn, J. H., Ekin, D., Alia-Klein, N., & Goldstein, R. Z. (2023). Common and distinct fronto-striatal volumetric changes in heroin and cocaine use disorders. *Brain*, *146*(4), 1662-1671. https://doi.org/10.1093/brain/awac366

Cheng, Y. C., Ryan, K. A., Qadwai, S. A., Shah, J., Sparks, M. J., Wozniak, M. A., Stern, B. J., Phipps, M. S., Cronin, C. A., Magder, L. S., Cole, J. W., & Kittner, S. J. (2016). Cocaine Use and Risk of Ischemic Stroke in Young Adults. *Stroke*, *47*(4), 918-922. https://doi.org/10.1161/strokeaha.115.011417

Collado-Torres, L., Klei, L., Liu, C., Kleinman, J. E., Hyde, T. M., Geschwind, D. H., Gandal, M. J., Devlin, B., & Weinberger, D. R. (2023). Comparison of gene expression in living and postmortem human brain. *medRxiv*. https://doi.org/10.1101/2023.11.08.23298172

Connolly, C. G., Bell, R. P., Foxe, J. J., & Garavan, H. (2013). Dissociated grey matter changes with prolonged addiction and extended abstinence in cocaine users. *PLoS One*, *8*(3), e59645. https://doi.org/10.1371/journal.pone.0059645

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938-2940. https://doi.org/10.1093/bioinformatics/btx364

Cornish, J. L., & Kalivas, P. W. (2000). Glutamate transmission in the nucleus accumbens mediates relapse in cocaine addiction. *J Neurosci*, *20*(15), Rc89. https://doi.org/10.1523/JNEUROSCI.20-15-j0006.2000

Correia, C., Romieu, P., Olmstead, M. C., & Befort, K. (2020). Can cocaine-induced neuroinflammation explain maladaptive cocaine-associated memories? *Neurosci Biobehav Rev*, *111*, 69-83. https://doi.org/10.1016/j.neubiorev.2020.01.001

Coviello, D. M., Alterman, A. I., Rutherford, M. J., Cacciola, J. S., McKay, J. R., & Zanis, D. A. (2001). The effectiveness of two intensities of psychosocial treatment for cocaine dependence. *Drug and Alcohol Dependence*, *61*(2), 145-154. https://doi.org/https://doi.org/10.1016/S0376-8716(00)00136-8

Cox, J., & Witten, I. B. (2019). Striatal circuits for reward learning and decision-making. *Nat Rev Neurosci*, *20*(8), 482-494. https://doi.org/10.1038/s41583-019-0189-2

Ctortecka, C., Clark, N. M., Boyle, B. W., Seth, A., Mani, D. R., Udeshi, N. D., & Carr, S. A. (2024). Automated single-cell proteomics providing sufficient proteome depth to study complex biology beyond cell type classifications. *Nature Communications*, *15*(1), 5707. https://doi.org/10.1038/s41467-024-49651-w

Cunha-Oliveira, T., Rego, A. C., Cardoso, S. M., Borges, F., Swerdlow, R. H., Macedo, T., & de Oliveira, C. R. (2006). Mitochondrial dysfunction and caspase activation in rat cortical neurons treated with cocaine or amphetamine. *Brain Res*, *1089*(1), 44-54. https://doi.org/10.1016/j.brainres.2006.03.061

Cunha-Oliveira, T., Silva, L., Silva, A. M., Moreno, A. J., Oliveira, C. R., & Santos, M. S. (2013). Mitochondrial complex I dysfunction induced by cocaine and cocaine plus morphine in brain and liver mitochondria. *Toxicology Letters*, *219*(3), 298-306. https://doi.org/https://doi.org/10.1016/j.toxlet.2013.03.025

Daley, D. C. (2013). Family and social aspects of substance use disorders and treatment. *J Food Drug Anal*, *21*(4), S73-s76. https://doi.org/10.1016/j.jfda.2013.09.038

Davies, T. H., Ning, Y. M., & Sánchez, E. R. (2002). A new first step in activation of steroid receptors: hormone-induced switching of FKBP51 and FKBP52 immunophilins. *J Biol Chem*, *277*(7), 4597-4600. https://doi.org/10.1074/jbc.C100531200

Davis, J. L., Kennedy, C., McMahon, C. L., Keegan, L., Clerkin, S., Treacy, N. J., Hoban, A. E., Kelly, Y., Brougham, D. F., Crean, J., & Murphy, K. J. (2025). Cocaine perturbs neurodevelopment and increases neuroinflammation in a prenatal cerebral organoid model. *Translational Psychiatry*, *15*(1), 94. https://doi.org/10.1038/s41398-025-03315-5

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol*, *11*(4), e1004219. https://doi.org/10.1371/journal.pcbi.1004219

Deak, J., Zhou, H., Levey, D., Galimberti, M., Farajzadeh, L., Hougaard, D., Jennings, M., Davis, L., Sanchez-Roige, S., Demontis, D., Børglum, A., Kranzler, H., Gaziano, M., Stein, M., & Gelernter, J. (2023). 79. A LARGE-SCALE GENOME-WIDE ASSOCIATION STUDY OF COCAINE USE DISORDER. *European Neuropsychopharmacology*, *75*, S99. https://doi.org/https://doi.org/10.1016/j.euroneuro.2023.08.184

Deak, J. D., & Johnson, E. C. (2021). Genetics of substance use disorders: a review. *Psychol Med*, *51*(13), 2189-2200. https://doi.org/10.1017/s0033291721000969

Dedic, N., Pöhlmann, M. L., Richter, J. S., Mehta, D., Czamara, D., Metzger, M. W., Dine, J., Bedenk, B. T., Hartmann, J., Wagner, K. V., Jurik, A., Almli, L. M., Lori, A., Moosmang, S., Hofmann, F., Wotjak, C. T., Rammes, G., Eder, M., Chen, A., . . . Deussing, J. M. (2018). Cross-disorder risk gene CACNA1C differentially modulates susceptibility to psychiatric disorders during development and adulthood. *Mol Psychiatry*, *23*(3), 533-543. https://doi.org/10.1038/mp.2017.133

Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., Mantilla-Herrara, A., Whiteford, H., Leung, J., Naghavi, M., Griswold, M., Rehm, J., Hall, W., Sartorius, B., Scott, J., Vollset, S. E., Knudsen, A. K., Haro, J. M., Patton, G., Kopec, J., . . . Vos, T. (2018). The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry*, *5*(12), 987-1012. https://doi.org/10.1016/S2215-0366(18)30337-7

DeMars, K. M., Ross, M. R., Starr, A., & McIntyre, J. C. (2023). Neuronal primary cilia integrate peripheral signals with metabolic drives. *Front Physiol*, *14*, 1150232. https://doi.org/10.3389/fphys.2023.1150232

Deroche-Gamonet, V., Belin, D., & Piazza, P. V. (2004). Evidence for addiction-like behavior in the rat. *Science*, *305*(5686), 1014-1017. https://doi.org/10.1126/science.1099020

Deroche-Gamonet, V., Sillaber, I., Aouizerate, B., Izawa, R., Jaber, M., Ghozland, S., Kellendonk, C., Le Moal, M., Spanagel, R., Schütz, G., Tronche, F., & Piazza, P. V. (2003). The glucocorticoid receptor as a potential target to reduce cocaine abuse. *J Neurosci*, *23*(11), 4785-4790. https://doi.org/10.1523/jneurosci.23-11-04785.2003

Devi, B. G., & Chan, A. W. (1997). Impairment of mitochondrial respiration and electron transport chain enzymes during cocaine-induced hepatic injury. *Life Sci*, *60*(11), 849-855. https://doi.org/10.1016/s0024-3205(97)00013-1

Dimitrakopoulos, C., Hindupur, S. K., Colombi, M., Liko, D., Ng, C. K. Y., Piscuoglio, S., Behr, J., Moore, A. L., Singer, J., Ruscheweyh, H. J., Matter, M. S., Mossmann, D., Terracciano, L. M., Hall, M. N., & Beerenwinkel, N. (2021). Multi-omics data integration reveals novel drug targets in

hepatocellular carcinoma. *BMC Genomics*, *22*(1), 592. https://doi.org/10.1186/s12864-021-07876-9

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15-21. https://doi.org/10.1093/bioinformatics/bts635

Doncheva, N. T., Morris, J. H., Gorodkin, J., & Jensen, L. J. (2019). Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res*, *18*(2), 623-632. https://doi.org/10.1021/acs.jproteome.8b00702

Du, C., Park, K., Hua, Y., Liu, Y., Volkow, N. D., & Pan, Y. (2024). Astrocytes modulate cerebral blood flow and neuronal response to cocaine in prefrontal cortex. *Molecular Psychiatry*, *29*(3), 820-834. https://doi.org/10.1038/s41380-023-02373-9

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, *11*(1), 587. https://doi.org/10.1186/1471-2105-11-587

Emmerich, C. H., Gamboa, L. M., Hofmann, M. C. J., Bonin-Andresen, M., Arbach, O., Schendel, P., Gerlach, B., Hempel, K., Bespalov, A., Dirnagl, U., & Parnham, M. J. (2021). Improving target assessment in biomedical research: the GOT-IT recommendations. *Nature Reviews Drug Discovery*, *20*(1), 64-81. https://doi.org/10.1038/s41573-020-0087-3

Engeli, E. J. E., Zoelch, N., Hock, A., Nordt, C., Hulka, L. M., Kirschner, M., Scheidegger, M., Esposito, F., Baumgartner, M. R., Henning, A., Seifritz, E., Quednow, B. B., & Herdener, M. (2021). Impaired glutamate homeostasis in the nucleus accumbens in human cocaine addiction. *Mol Psychiatry*, *26*(9), 5277-5285. https://doi.org/10.1038/s41380-020-0828-z

Ersche, K. D., Barnes, A., Jones, P. S., Morein-Zamir, S., Robbins, T. W., & Bullmore, E. T. (2011). Abnormal structure of frontostriatal brain systems is associated with aspects of impulsivity and compulsivity in cocaine dependence. *Brain*, *134*(Pt 7), 2013-2024. https://doi.org/10.1093/brain/awr138

Ersche, K. D., Jones, P. S., Williams, G. B., Robbins, T. W., & Bullmore, E. T. (2013). Cocaine dependence: a fast-track for brain ageing? *Mol Psychiatry*, *18*(2), 134-135. https://doi.org/10.1038/mp.2012.31

Ersche, K. D., Stochl, J., Woodward, J. M., & Fletcher, P. C. (2013). The skinny on cocaine: insights into eating behavior and body weight in cocaine-dependent men. *Appetite*, *71*, 75-80. https://doi.org/10.1016/j.appet.2013.07.011

European Monitoring Centre for Drugs and Drug Addiction. (2022). European Drug Report 2022: Trends and Developments. Luxembourg: Publications Office of the European Union.

Evan, N. G., Robert, A. W., David, A. B., Amanda, L. E., Jonathan, E. H., Jayme, R. M., Mykel, A. R., Oliver, V., Daniel, S. W., John, R. M., & Paul, J. G. (2013). Corticosterone Acts in the Nucleus Accumbens to Enhance Dopamine Signaling and Potentiate Reinstatement of Cocaine Seeking. *The Journal of Neuroscience*, *33*(29), 11800. https://doi.org/10.1523/JNEUROSCI.1969-13.2013

Evans, S. M., & Foltin, R. W. (2010). Does the response to cocaine differ as a function of sex or hormonal status in human and non-human primates? *Horm Behav*, *58*(1), 13-21. https://doi.org/10.1016/j.yhbeh.2009.08.010

Everett, T., Ten Eyck, T. W., Wu, C. H., Shelowitz, A. L., Stansbury, S. M., Firek, A., Setlow, B., & McIntyre, J. C. (2024). Cilia loss on distinct neuron populations differentially alters cocaine-induced locomotion and reward. *J Psychopharmacol*, *38*(2), 200-212. https://doi.org/10.1177/02698811231219058

Everitt, B. J., & Robbins, T. W. (2013). From the ventral to the dorsal striatum: Devolving views of their roles in drug addiction. *Neuroscience & Biobehavioral Reviews*, *37*(9, Part A), 1946-1954. https://doi.org/https://doi.org/10.1016/j.neubiorev.2013.02.010

Feil, R., & Fraga, M. F. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, *13*(2), 97-109. https://doi.org/10.1038/nrg3142

Feng, J., Shao, N., Szulwach, K. E., Vialou, V., Huynh, J., Zhong, C., Le, T., Ferguson, D., Cahill, M. E., Li, Y., Koo, J. W., Ribeiro, E., Labonte, B., Laitman, B. M., Estey, D., Stockman, V., Kennedy, P., Couroussé, T., Mensah, I., . . . Nestler, E. J. (2015). Role of Tet1 and 5-hydroxymethylcytosine in cocaine action. *Nat Neurosci*, *18*(4), 536-544. https://doi.org/10.1038/nn.3976

Fernàndez-Castillo, N., Cabana-Domínguez, J., Corominas, R., & Cormand, B. (2022). Molecular genetics of cocaine use disorders in humans. *Molecular Psychiatry*, *27*(1), 624-639. https://doi.org/10.1038/s41380-021-01256-1

Ford, J. D., Gelernter, J., DeVoe, J. S., Zhang, W., Weiss, R. D., Brady, K., Farrer, L., & Kranzler, H. R. (2009). Association of psychiatric and substance use disorder comorbidity with cocaine dependence severity and treatment utilization in cocaine-dependent individuals. *Drug Alcohol Depend*, *99*(1-3), 193-203. https://doi.org/10.1016/j.drugalcdep.2008.07.004

Friedman, M. J., Huber, B. R., Brady, C. B., Ursano, R. J., Benedek, D. M., Kowall, N. W., & McKee, A. C. (2017). VA's National PTSD Brain Bank: a National Resource for Research. *Curr Psychiatry Rep*, *19*(10), 73. https://doi.org/10.1007/s11920-017-0822-6

Frye, M. A., Hinton, D. J., Karpyak, V. M., Biernacka, J. M., Gunderson, L. J., Geske, J., Feeder, S. E., Choi, D. S., & Port, J. D. (2016). Elevated Glutamate Levels in the Left Dorsolateral Prefrontal Cortex Are Associated with Higher Cravings for Alcohol. *Alcohol Clin Exp Res*, *40*(8), 1609-1616. https://doi.org/10.1111/acer.13131

Gao, P., Limpens, J. H. W., Spijker, S., Vanderschuren, L. J. M. J., & Voorn, P. (2017). Stable immediate early gene expression patterns in medial prefrontal cortex and striatum after long-term cocaine self-administration. *Addiction Biology*, *22*(2), 354-368. https://doi.org/https://doi.org/10.1111/adb.12330

Garavan, H., Pankiewicz, J., Bloom, A., Cho, J. K., Sperry, L., Ross, T. J., Salmeron, B. J., Risinger, R., Kelley, D., & Stein, E. A. (2000). Cue-induced cocaine craving: neuroanatomical specificity for drug users and drug stimuli. *Am J Psychiatry*, *157*(11), 1789-1798. https://doi.org/10.1176/appi.ajp.157.11.1789

Garcia-Gonzalo, F. R., Phua, S. C., Roberson, E. C., Garcia, G., 3rd, Abedin, M., Schurmans, S., Inoue, T., & Reiter, J. F. (2015). Phosphoinositides Regulate Ciliary Protein Trafficking to Modulate Hedgehog Signaling. *Dev Cell*, *34*(4), 400-409. https://doi.org/10.1016/j.devcel.2015.08.001

Garma, L. D., Harder, L., Barba-Reyes, J. M., Marco Salas, S., Díez-Salguero, M., Nilsson, M., Serrano-Pozo, A., Hyman, B. T., & Muñoz-Manchado, A. B. (2024). Interneuron diversity in the human dorsal striatum. *Nature Communications*, *15*(1), 6164. https://doi.org/10.1038/s41467-024-50414-w

Gaudreault, P. O., King, S. G., Malaker, P., Alia-Klein, N., & Goldstein, R. Z. (2023). Whole-brain white matter abnormalities in human cocaine and heroin use disorders: association with craving, recency, and cumulative use. *Mol Psychiatry*, *28*(2), 780-791. https://doi.org/10.1038/s41380-022-01833-y

Gelernter, J., Sherva, R., Koesterer, R., Almasy, L., Zhao, H., Kranzler, H. R., & Farrer, L. (2014). Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol Psychiatry*, *19*(6), 717-723. https://doi.org/10.1038/mp.2013.99

Gibney, E. R., & Nolan, C. M. (2010). Epigenetics and gene expression. *Heredity*, *105*(1), 4-13. https://doi.org/10.1038/hdy.2010.54

Giordano, T. P., Tropea, T. F., Satpute, S. S., Sinnegger-Brauns, M. J., Striessnig, J., Kosofsky, B. E., & Rajadhyaksha, A. M. (2010). Molecular switch from L-type Ca v 1.3 to Ca v 1.2 Ca2+ channel signaling underlies long-term psychostimulant-induced behavioral and molecular plasticity. *J Neurosci*, *30*(50), 17051-17062. https://doi.org/10.1523/jneurosci.2255-10.2010

Girgenti, M. J., Wang, J., Ji, D., Cruz, D. A., Stein, M. B., Gelernter, J., Young, K. A., Huber, B. R., Williamson, D. E., Friedman, M. J., Krystal, J. H., Zhao, H., & Duman, R. S. (2021). Transcriptomic organization of the human

brain in post-traumatic stress disorder. *Nat Neurosci*, *24*(1), 24-33. https://doi.org/10.1038/s41593-020-00748-7

Gokce, O., Stanley, Geoffrey M., Treutlein, B., Neff, Norma F., Camp, J. G., Malenka, Robert C., Rothwell, Patrick E., Fuccillo, Marc V., Südhof, Thomas C., & Quake, Stephen R. (2016). Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *Cell Reports*, *16*(4), 1126-1137. https://doi.org/https://doi.org/10.1016/j.celrep.2016.06.059

Gold, M. S. (1993). The History of Cocaine. In *Cocaine* (pp. 11-35). Springer US. https://doi.org/10.1007/978-1-4684-6033-9_2

Goldstein, R. Z., & Volkow, N. D. (2002). Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *Am J Psychiatry*, *159*(10), 1642-1652. https://doi.org/10.1176/appi.ajp.159.10.1642

Goldstein, R. Z., & Volkow, N. D. (2011). Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nature Reviews Neuroscience*, *12*(11), 652-669. https://doi.org/10.1038/nrn3119

Gomez-Mancilla, B., Dürsteler, K. M., Vogel, M., Herdener, M., Torrens, M., Gálvez, B. P., Gual, A., Corral, R. M., Kuper, E. I., Mosca, D. L., Dumitras, S., Pezous, N., Berkheimer, M., Walker, E., Gasparini, F., Cha, J.-H., & Dolmetsch, R. (2025). Mavoglurant reduces cocaine use in patients with cocaine use disorder in a phase 2 clinical trial. *Science Translational Medicine*, *17*(792), eadi4505. https://doi.org/doi:10.1126/scitranslmed.adi4505

Green, N., Gao, H., Chu, X., Yuan, Q., McGuire, P., Lai, D., Jiang, G., Xuei, X., Reiter, J. L., Stevens, J., Sutherland, G. T., Goate, A. M., Pang, Z. P., Slesinger, P. A., Hart, R. P., Tischfield, J. A., Agrawal, A., Wang, Y., Duren, Z., . . . Liu, Y. (2024). Integrated Single-Cell Multiomic Profiling of Caudate Nucleus Suggests Key Mechanisms in Alcohol Use Disorder. *bioRxiv*. https://doi.org/10.1101/2024.08.02.606355

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847-2849. https://doi.org/10.1093/bioinformatics/btw313

Guardia, C. M., De Pace, R., Mattera, R., & Bonifacino, J. S. (2018). Neuronal functions of adaptor complexes involved in protein sorting. *Curr Opin Neurobiol*, *51*, 103-110. https://doi.org/10.1016/j.conb.2018.02.021

Hall, F. S., Li, X. F., Randall-Thompson, J., Sora, I., Murphy, D. L., Lesch, K. P., Caron, M., & Uhl, G. R. (2009). Cocaine-conditioned locomotion in dopamine transporter, norepinephrine transporter and 5-HT transporter knockout mice. *Neuroscience*, *162*(4), 870-880. https://doi.org/https://doi.org/10.1016/j.neuroscience.2009.05.058

Hámor, P. U., Edelmann, M. J., Gobin, C., & Schwendt, M. (2020). Long-term Changes in the Central Amygdala Proteome in Rats with a History of Chronic Cocaine Self-administration. *Neuroscience*, *443*, 93-109. https://doi.org/10.1016/j.neuroscience.2020.06.011

Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., & Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology*, *42*(2), 293-304. https://doi.org/10.1038/s41587-023-01767-y

Hardee, I., Soldatos, A., Davids, M., Vilboux, T., Toro, C., David, K. L., Ferreira, C. R., Nehrebecky, M., Snow, J., Thurm, A., Heller, T., Macnamara, E. F., Gunay-Aygun, M., Zein, W. M., Gahl, W. A., & Malicdan, M. C. V. (2017). Defective ciliogenesis in INPP5E-related Joubert syndrome. *Am J Med Genet A*, *173*(12), 3231-3237. https://doi.org/10.1002/ajmg.a.38376

Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biology*, *18*(1), 83. https://doi.org/10.1186/s13059-017-1215-1

He, L., Davila-Velderrain, J., Sumida, T. S., Hafler, D. A., Kellis, M., & Kulminski, A. M. (2021). NEBULA is a fast negative binomial mixed model for differential or co-expression analysis of large-scale multi-subject single-cell data. *Communications Biology*, *4*(1), 629. https://doi.org/10.1038/s42003-021-02146-6

Heinz, A., Kiefer, F., Smolka, M. N., Endrass, T., Beste, C., Beck, A., Liu, S., Genauck, A., Romund, L., Banaschewski, T., Bermpohl, F., Deserno, L., Dolan, R. J., Durstewitz, D., Ebner-Priemer, U., Flor, H., Hansson, A. C., Heim, C., Hermann, D., . . . Spanagel, R. (2020). Addiction Research Consortium: Losing and regaining control over drug intake (ReCoDe)-From trajectories to mechanisms and interventions. *Addict Biol*, *25*(2), e12866. https://doi.org/10.1111/adb.12866

Hester, R., & Garavan, H. (2004). Executive Dysfunction in Cocaine Addiction: Evidence for Discordant Frontal, Cingulate, and Cerebellar Activity. *The Journal of Neuroscience*, *24*(49), 11017. https://doi.org/10.1523/JNEUROSCI.3321-04.2004

Higgins, S. T., Wong, C. J., Badger, G. J., Ogden, D. E., & Dantona, R. L. (2000). Contingent reinforcement increases cocaine abstinence during outpatient treatment and 1 year of follow-up. *J Consult Clin Psychol*, *68*(1), 64-72. https://doi.org/10.1037//0022-006x.68.1.64

Hillen, A. E. J., Burbach, J. P. H., & Hol, E. M. (2018). Cell adhesion and matricellular support by astrocytes of the tripartite synapse. *Progress in Neurobiology*, *165-167*, 66-86. https://doi.org/https://doi.org/10.1016/j.pneurobio.2018.02.002

Hoffman, G. E., & Schadt, E. E. (2016). variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics*, *17*(1), 483. https://doi.org/10.1186/s12859-016-1323-z

Hope, B., Kosofsky, B., Hyman, S. E., & Nestler, E. J. (1992). Regulation of immediate early gene expression and AP-1 binding in the rat nucleus accumbens by chronic cocaine. *Proc Natl Acad Sci U S A*, *89*(13), 5764-5768. https://doi.org/10.1073/pnas.89.13.5764

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), 3156. https://doi.org/10.1186/gb-2013-14-10-r115

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., Wiencke, J. K., & Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, *13*, 86. https://doi.org/10.1186/1471-2105-13-86

Howell, L. L., Votaw, J. R., Goodman, M. M., & Lindsey, K. P. (2010). Cortical activation during cocaine use and extinction in rhesus monkeys. *Psychopharmacology (Berl)*, *208*(2), 191-199. https://doi.org/10.1007/s00213-009-1720-3

Hu, Y., Salmeron, B. J., Gu, H., Stein, E. A., & Yang, Y. (2015). Impaired functional connectivity within and between frontostriatal circuits and its association with compulsive drug use and trait impulsivity in cocaine addiction. *JAMA Psychiatry*, *72*(6), 584-592. https://doi.org/10.1001/jamapsychiatry.2015.1

Hu, Y., Salmeron, B. J., Gu, H., Stein, E. A., & Yang, Y. (2015). Impaired Functional Connectivity Within and Between Frontostriatal Circuits and Its Association With Compulsive Drug Use and Trait Impulsivity in Cocaine Addiction. *JAMA Psychiatry*, *72*(6), 584-592. https://doi.org/10.1001/jamapsychiatry.2015.1

Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A., & Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, *18 Suppl 1*, S96-104. https://doi.org/10.1093/bioinformatics/18.suppl_1.s96

Huggett, S. B., Ikeda, A. S., McGeary, J. E., Kaun, K. R., & Palmer, R. H. C. (2022). Opioid Use Disorder and Alternative mRNA Splicing in Reward Circuitry. *Genes (Basel)*, *13*(6). https://doi.org/10.3390/genes13061045

Huggett, S. B., Ikeda, A. S., Yuan, Q., Benca-Bachman, C. E., & Palmer, R. H. C. (2023). Genome- and transcriptome-wide splicing associations with alcohol use disorder. *Sci Rep*, *13*(1), 3950. https://doi.org/10.1038/s41598-023-30926-z

Huggett, S. B., & Stallings, M. C. (2020). Cocaine'omics: Genome-wide and transcriptome-wide analyses provide biological insight into cocaine use

and dependence. *Addict Biol*, *25*(2), e12719. https://doi.org/10.1111/adb.12719

Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., & Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol*, *10*(10), 757. https://doi.org/10.15252/msb.20145625

Hughes, C. S., Moggridge, S., Müller, T., Sorensen, P. H., Morin, G. B., & Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature Protocols*, *14*(1), 68-85. https://doi.org/10.1038/s41596-018-0082-x

An integrated encyclopedia of DNA elements in the human genome. (2012). *Nature*, *489*(7414), 57-74. https://doi.org/10.1038/nature11247

Jaffe, A. E., & Kaminsky, Z. (2017). FlowSorted. *Blood. 450k: Illumina HumanMethylation data on sorted blood cell populations. R package version*, *1*(0).

Jedema, H. P., Song, X., Aizenstein, H. J., Bonner, A. R., Stein, E. A., Yang, Y., & Bradberry, C. W. (2021). Long-Term Cocaine Self-administration Produces Structural Brain Changes That Correlate With Altered Cognition. *Biol Psychiatry*, *89*(4), 376-385. https://doi.org/10.1016/j.biopsych.2020.08.008

Jin, S., Guerrero-Juarez, C. F., Zhang, L., Chang, I., Ramos, R., Kuan, C.-H., Myung, P., Plikus, M. V., & Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. *Nature Communications*, *12*(1), 1088. https://doi.org/10.1038/s41467-021-21246-9

Joehanes, R., Just, A. C., Marioni, R. E., Pilling, L. C., Reynolds, L. M., Mandaviya, P. R., Guan, W., Xu, T., Elks, C. E., Aslibekyan, S., Moreno-Macias, H., Smith, J. A., Brody, J. A., Dhingra, R., Yousefi, P., Pankow, J. S., Kunze, S., Shah, S. H., McRae, A. F., . . . London, S. J. (2016). Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet*, *9*(5), 436-447. https://doi.org/10.1161/circgenetics.116.001506

John, R. M., David, S., & Nick, E. G. (1998). Corticosterone Facilitates the Acquisition of Cocaine Self-Administration in Rats: Opposite Effects of the Type II Glucocorticoid Receptor Agonist Dexamethasone. *Journal of Pharmacology and Experimental Therapeutics*, *287*(1), 72. http://jpet.aspetjournals.org/content/287/1/72.abstract

Johnson, B. A., Ait-Daoud, N., Wang, X.-Q., Penberthy, J. K., Javors, M. A., Seneviratne, C., & Liu, L. (2013). Topiramate for the Treatment of Cocaine Addiction: A Randomized Clinical Trial. *JAMA Psychiatry*, *70*(12), 1338-1346. https://doi.org/10.1001/jamapsychiatry.2013.2295

Johnson, E. O., Fisher, H. S., Sullivan, K. A., Corradin, O., Sanchez-Roige, S., Gaddis, N. C., Sami, Y. N., Townsend, A., Teixeira Prates, E., Pavicic, M., Kruse, P., Chesler, E. J., Palmer, A. A., Troiani, V., Bubier, J. A., Jacobson,

D. A., & Maher, B. S. (2024). An emerging multi-omic understanding of the genetics of opioid addiction. *The Journal of Clinical Investigation*, *134*(20). https://doi.org/10.1172/JCI172886

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews Genetics*, *16*(7), 421-433. https://doi.org/10.1038/nrg3965

Kalivas, P. W., Volkow, N., & Seamans, J. (2005). Unmanageable Motivation in Addiction: A Pathology in Prefrontal-Accumbens Glutamate Transmission. *Neuron*, *45*(5), 647-650. https://doi.org/https://doi.org/10.1016/j.neuron.2005.02.005

Kampman, K. M. (2019). The treatment of cocaine use disorder. *Sci Adv*, *5*(10), eaax1532. https://doi.org/10.1126/sciadv.aax1532

Kampman, K. M., Dackis, C., Pettinati, H. M., Lynch, K. G., Sparkman, T., & O'Brien, C. P. (2011). A double-blind, placebo-controlled pilot trial of acamprosate for the treatment of cocaine dependence. *Addict Behav*, *36*(3), 217-221. https://doi.org/10.1016/j.addbeh.2010.11.003

Kampman, K. M., Pettinati, H. M., Lynch, K. G., Spratt, K., Wierzbicki, M. R., & O'Brien, C. P. (2013). A double-blind, placebo-controlled trial of topiramate for the treatment of comorbid cocaine and alcohol dependence. *Drug Alcohol Depend*, *133*(1), 94-99. https://doi.org/10.1016/j.drugalcdep.2013.05.026

Kan, Z., Wen, J., Bonato, V., Webster, J., Yang, W., Ivanov, V., Kim, K. H., Roh, W., Liu, C., Mu, X. J., Lapira-Miller, J., Oyer, J., VanArsdale, T., Rejto, P. A., & Bienkowska, J. (2025). Real-world clinical multi-omics analyses reveal bifurcation of ER-independent and ER-dependent drug resistance to CDK4/6 inhibitors. *Nature Communications*, *16*(1), 932. https://doi.org/10.1038/s41467-025-55914-x

Kasanetz, F., Deroche-Gamonet, V., Berson, N., Balado, E., Lafourcade, M., Manzoni, O., & Piazza, P. V. (2010). Transition to Addiction Is Associated with a Persistent Impairment in Synaptic Plasticity. *Science*, *328*(5986), 1709-1712. https://doi.org/doi:10.1126/science.1187801

Kendler, K. S., Karkowski, L. M., Neale, M. C., & Prescott, C. A. (2000). Illicit Psychoactive Substance Use, Heavy Use, Abuse, and Dependence in a US Population-Based Sample of Male Twins. *Archives of General Psychiatry*, *57*(3), 261-269. https://doi.org/10.1001/archpsyc.57.3.261

Kendler, K. S., Myers, J., & Prescott, C. A. (2007). Specificity of genetic and environmental risk factors for symptoms of cannabis, cocaine, alcohol, caffeine, and nicotine dependence. *Arch Gen Psychiatry*, *64*(11), 1313-1320. https://doi.org/10.1001/archpsyc.64.11.1313

King, S. G., Gaudreault, P. O., Malaker, P., Kim, J. W., Alia-Klein, N., Xu, J., & Goldstein, R. Z. (2022). Prefrontal-habenular microstructural impairments

in human cocaine and heroin addiction. *Neuron*, *110*(22), 3820-3832.e3824. https://doi.org/10.1016/j.neuron.2022.09.011

Kiraly, D. D., Nemirovsky, N. E., LaRese, T. P., Tomek, S. E., Yahn, S. L., Olive, M. F., Eipper, B. A., & Mains, R. E. (2013). Constitutive knockout of kalirin-7 leads to increased rates of cocaine self-administration. *Mol Pharmacol*, *84*(4), 582-590. https://doi.org/10.1124/mol.113.087106

Knackstedt, L. A., Melendez, R. I., & Kalivas, P. W. (2010). Ceftriaxone restores glutamate homeostasis and prevents relapse to cocaine seeking. *Biol Psychiatry*, *67*(1), 81-84. https://doi.org/10.1016/j.biopsych.2009.07.018

Knouse, M. C., & Briand, L. A. (2021). Behavioral sex differences in cocaine and opioid use disorders: The role of gonadal hormones. *Neuroscience & Biobehavioral Reviews*, *128*, 358-366. https://doi.org/https://doi.org/10.1016/j.neubiorev.2021.06.038

Kodam, P., Sai Swaroop, R., Pradhan, S. S., Sivaramakrishnan, V., & Vadrevu, R. (2023). Integrated multi-omics analysis of Alzheimer's disease shows molecular signatures associated with disease progression and potential therapeutic targets. *Sci Rep*, *13*(1), 3695. https://doi.org/10.1038/s41598-023-30892-6

Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., & Nesvizhskii, A. I. (2017). MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nature Methods*, *14*(5), 513-520. https://doi.org/10.1038/nmeth.4256

Koob, G. F. (2008). A Role for Brain Stress Systems in Addiction. *Neuron*, *59*(1), 11-34. https://doi.org/https://doi.org/10.1016/j.neuron.2008.06.012

Koob, G. F., Buck, C. L., Cohen, A., Edwards, S., Park, P. E., Schlosburg, J. E., Schmeichel, B., Vendruscolo, L. F., Wade, C. L., Whitfield, T. W., & George, O. (2014). Addiction as a stress surfeit disorder. *Neuropharmacology*, *76*, 370-382. https://doi.org/https://doi.org/10.1016/j.neuropharm.2013.05.024

Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of Addiction. *Neuropsychopharmacology*, *35*(1), 217-238. https://doi.org/10.1038/npp.2009.110

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry*, *3*(8), 760-773. https://doi.org/10.1016/s2215-0366(16)00104-8

Kotlinska, J., Pachuta, A., & Silberring, J. (2008). Neuropeptide FF (NPFF) reduces the expression of cocaine-induced conditioned place preference and cocaine-induced sensitization in animals. *Peptides*, *29*(6), 933-939. https://doi.org/10.1016/j.peptides.2008.01.008

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *J Exp Psychol Gen*, *142*(2), 573-603. https://doi.org/10.1037/a0029146

Kruyer, A., & Scofield, M. D. (2021). Astrocytes in Addictive Disorders. *Adv Neurobiol*, *26*, 231-254. https://doi.org/10.1007/978-3-030-77375-5_10

Kurtenbach, S., & Harbour, J. W. (2019). SparK: A Publication-quality NGS Visualization Tool. *bioRxiv*, 845529. https://doi.org/10.1101/845529

Lange, R. A., Cigarroa, R. G., Yancy, C. W., Willard, J. E., Popma, J. J., Sills, M. N., McBride, W., Kim, A. S., & Hillis, L. D. (1989). Cocaine-Induced Coronary-Artery Vasoconstriction. *New England Journal of Medicine*, *321*(23), 1557-1562. https://doi.org/doi:10.1056/NEJM198912073212301

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Lee, D., Takayama, S., & Goldberg, A. L. (2018). ZFAND5/ZNF216 is an activator of the 26S proteasome that stimulates overall protein degradation. *Proceedings of the National Academy of Sciences*, *115*(41), E9550-E9559. https://doi.org/10.1073/pnas.1809934115

Lehne, B., Drong, A. W., Loh, M., Zhang, W., Scott, W. R., Tan, S. T., Afzal, U., Scott, J., Jarvelin, M. R., Elliott, P., McCarthy, M. I., Kooner, J. S., & Chambers, J. C. (2015). A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol*, *16*(1), 37. https://doi.org/10.1186/s13059-015-0600-x

Lent, S., Cardenas, A., Rifas-Shiman, S. L., Perron, P., Bouchard, L., Liu, C. T., Hivert, M. F., & Dupuis, J. (2021). Detecting differentially methylated regions with multiple distinct associations. *Epigenomics*, *13*(6), 451-464. https://doi.org/10.2217/epi-2020-0344

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*, *10*(4), 573-591. https://doi.org/10.18632/aging.101414

Li, C. S., Huang, C., Yan, P., Bhagwagar, Z., Milivojevic, V., & Sinha, R. (2008). Neural correlates of impulse control during stop signal inhibition in cocaine-dependent men. *Neuropsychopharmacology*, *33*(8), 1798-1806. https://doi.org/10.1038/sj.npp.1301568

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352

Li, M., Xu, P., Xu, Y., Teng, H., Tian, W., Du, Q., & Zhao, M. (2017). Dynamic Expression Changes in the Transcriptome of the Prefrontal Cortex after Repeated Exposure to Cocaine in Mice. *Front Pharmacol*, *8*, 142. https://doi.org/10.3389/fphar.2017.00142

Li, R., Reiter, J. L., Chen, A. B., Chen, S. X., Foroud, T., Edenberg, H. J., Lai, D., & Liu, Y. (2023). RNA alternative splicing impacts the risk for alcohol use disorder. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-023-02111-1

Li, Y. I., Knowles, D. A., Humphrey, J., Barbeira, A. N., Dickinson, S. P., Im, H. K., & Pritchard, J. K. (2018). Annotation-free quantification of RNA splicing using LeafCutter. *Nature Genetics*, *50*(1), 151-158. https://doi.org/10.1038/s41588-017-0004-9

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, *47*(8), e47-e47. https://doi.org/10.1093/nar/gkz114

Liharska, L. E., Park, Y. J., Ziafat, K., Wilkins, L., Silk, H., Linares, L. M., Vornholt, E., Sullivan, B., Cohen, V., Kota, P., Feng, C., Cheng, E., Moya, E., Thompson, R. C., Johnson, J. S., Rieder, M. K., Huang, J., Scarpa, J., Hashemi, A., . . . Charney, A. W. (2023). A study of gene expression in the living human brain. *medRxiv*. https://doi.org/10.1101/2023.04.21.23288916

Lin, Y. T., Yu, Y. L., Hong, W. C., Yeh, T. S., Chen, T. C., & Chen, J. C. (2017). NPFFR2 Activates the HPA Axis and Induces Anxiogenic Effects in Rodents. *Int J Mol Sci*, *18*(8). https://doi.org/10.3390/ijms18081810

Liu, A., Dai, Y., Mendez, E. F., Hu, R., Fries, G. R., Najera, K. E., Jiang, S., Meyer, T. D., Stertz, L., Jia, P., Walss-Bass, C., & Zhao, Z. (2021). Genome-Wide Correlation of DNA Methylation and Gene Expression in Postmortem Brain Tissues of Opioid Use Disorder Patients. *Int J Neuropsychopharmacol*, *24*(11), 879-891. https://doi.org/10.1093/ijnp/pyab043

Liu, Y., Guazzelli Williamson, V., Setlow, B., Cottler, L. B., & Knackstedt, L. A. (2018). The importance of considering polysubstance use: lessons from cocaine research. *Drug Alcohol Depend*, *192*, 16-28. https://doi.org/10.1016/j.drugalcdep.2018.07.025

Lohoff, F. W., Roy, A., Jung, J., Longley, M., Rosoff, D. B., Luo, A., O'Connell, E., Sorcher, J. L., Sun, H., Schwandt, M., Hodgkinson, C. A., Goldman, D., Momenan, R., McIntosh, A. M., Adams, M. J., Walker, R. M., Evans, K. L., Porteous, D., Smith, A. K., . . . Kaminsky, Z. A. (2021). Epigenome-wide association study and multi-tissue replication of individuals with alcohol use disorder: evidence for abnormal glucocorticoid signaling pathway gene regulation. *Molecular Psychiatry*, *26*(6), 2224-2237. https://doi.org/10.1038/s41380-020-0734-4

López-Pedrajas, R., Ramírez-Lamelas, D. T., Muriach, B., Sánchez-Villarejo, M. V., Almansa, I., Vidal-Gil, L., Romero, F. J., Barcia, J. M., & Muriach, M. (2015). Cocaine promotes oxidative stress and microglial-macrophage activation in rat cerebellum. *Front Cell Neurosci*, *9*, 279. https://doi.org/10.3389/fncel.2015.00279

Lopez-Quintero, C., Pérez de los Cobos, J., Hasin, D. S., Okuda, M., Wang, S., Grant, B. F., & Blanco, C. (2011). Probability and predictors of transition from first use to dependence on nicotine, alcohol, cannabis, and cocaine: results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Drug Alcohol Depend*, *115*(1-2), 120-130. https://doi.org/10.1016/j.drugalcdep.2010.11.004

López-Valverde, A., de Vicente, J., Martínez-Domínguez, L., & de Diego, R. G. (2014). Local anaesthesia through the action of cocaine, the oral mucosa and the Vienna group. *British Dental Journal*, *217*(1), 41-43. https://doi.org/10.1038/sj.bdj.2014.546

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lövkvist, C., Dodd, I. B., Sneppen, K., & Haerter, J. O. (2016). DNA methylation in human epigenomes depends on local topology of CpG sites. *Nucleic Acids Research*, *44*(11), 5123-5132. https://doi.org/10.1093/nar/gkw124

Luo, J., Li, L., Niu, M., Kong, D., Jiang, Y., Poudel, S., Shieh, A. W., Cheng, L., Giase, G., Grennan, K., White, K. P., Chen, C., Wang, S. H., Pinto, D., Wang, Y., Liu, C., Peng, J., & Wang, X. (2024). Genetic regulation of human brain proteome reveals proteins implicated in psychiatric disorders. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-024-02576-8

Lüscher, C., & Malenka, Robert C. (2011). Drug-Evoked Synaptic Plasticity in Addiction: From Molecular Changes to Circuit Remodeling. *Neuron*, *69*(4), 650-663. https://doi.org/https://doi.org/10.1016/j.neuron.2011.01.017

Lynch, W. J. (2008). Acquisition and maintenance of cocaine self-administration in adolescent rats: effects of sex and gonadal hormones. *Psychopharmacology*, *197*(2), 237-246. https://doi.org/10.1007/s00213-007-1028-0

Ma, X. M., Kiraly, D. D., Gaier, E. D., Wang, Y., Kim, E. J., Levine, E. S., Eipper, B. A., & Mains, R. E. (2008). Kalirin-7 is required for synaptic structure and function. *J Neurosci*, *28*(47), 12368-12382. https://doi.org/10.1523/jneurosci.4269-08.2008

Mai, J. K., Paxinos, G., & Voss, T. (2007). *Atlas of the Human Brain* (3 ed.). Academic Press.

Marasco, L. E., & Kornblihtt, A. R. (2023). The physiology of alternative splicing. *Nature Reviews Molecular Cell Biology*, *24*(4), 242-254. https://doi.org/10.1038/s41580-022-00545-z

Martinez, L. A., Lees, M. E., Ruskin, D. N., & Masino, S. A. (2019). A ketogenic diet diminishes behavioral responses to cocaine in young adult male and female rats. *Neuropharmacology*, *149*, 27-34. https://doi.org/10.1016/j.neuropharm.2019.02.001

Maude-Griffin, P. M., Hohenstein, J. M., Humfleet, G. L., Reilly, P. M., Tusel, D. J., & Hall, S. M. (1998). Superior efficacy of cognitive-behavioral therapy for urban crack cocaine abusers: main and matching effects. *J Consult Clin Psychol*, *66*(5), 832-837. https://doi.org/10.1037//0022-006x.66.5.832

McClung, C. A., & Nestler, E. J. (2008). Neuroplasticity mediated by altered gene expression. *Neuropsychopharmacology*, *33*(1), 3-17. https://doi.org/10.1038/sj.npp.1301544

McFarland, K., Lapish, C. C., & Kalivas, P. W. (2003). Prefrontal glutamate release into the core of the nucleus accumbens mediates cocaine-induced reinstatement of drug-seeking behavior. *J Neurosci*, *23*(8), 3531-3537. https://doi.org/10.1523/jneurosci.23-08-03531.2003

Mendez, E. F., Wei, H., Hu, R., Stertz, L., Fries, G. R., Wu, X., Najera, K. E., Monterey, M. D., Lincoln, C. M., Kim, J.-w., Moriel, K., Meyer, T. D., Selvaraj, S., Teixeira, A. L., Zhao, Z., Xu, J., Wu, J., & Walss-Bass, C. (2021). Angiogenic gene networks are dysregulated in opioid use disorder: evidence from multi-omics and imaging of postmortem human brain. *Molecular Psychiatry*, *26*(12), 7803-7812. https://doi.org/10.1038/s41380-021-01259-y

Mews, P., Cunningham, A. M., Scarpa, J., Ramakrishnan, A., Hicks, E. M., Bolnick, S., Garamszegi, S., Shen, L., Mash, D. C., & Nestler, E. J. (2023). Convergent abnormalities in striatal gene networks in human cocaine use disorder and mouse cocaine administration models. *Science Advances*, *9*(6), eadd8946. https://doi.org/doi:10.1126/sciadv.add8946

Michels, K. B., Binder, A. M., Dedeurwaerder, S., Epstein, C. B., Greally, J. M., Gut, I., Houseman, E. A., Izzi, B., Kelsey, K. T., Meissner, A., Milosavljevic, A., Siegmund, K. D., Bock, C., & Irizarry, R. A. (2013). Recommendations for the design and analysis of epigenome-wide association studies. *Nature Methods*, *10*(10), 949-955. https://doi.org/10.1038/nmeth.2632

Miller, J. L., & Grant, P. A. (2013). The role of DNA methylation and histone modifications in transcriptional regulation in humans. *Subcell Biochem*, *61*, 289-317. https://doi.org/10.1007/978-94-007-4525-4_13

Mittleman, M. A., Mintzer, D., Maclure, M., Tofler, G. H., Sherwood, J. B., & Muller, J. E. (1999). Triggering of Myocardial Infarction by Cocaine.

Circulation, *99*(21), 2737-2741. https://doi.org/doi:10.1161/01.CIR.99.21.2737

Moeller, S. J., Konova, A. B., Parvaz, M. A., Tomasi, D., Lane, R. D., Fort, C., & Goldstein, R. Z. (2014). Functional, structural, and emotional correlates of impaired insight in cocaine addiction. *JAMA Psychiatry*, *71*(1), 61-70. https://doi.org/10.1001/jamapsychiatry.2013.2833

Mooney, M. E., Herin, D. V., Schmitz, J. M., Moukaddam, N., Green, C. E., & Grabowski, J. (2009). Effects of oral methamphetamine on cocaine use: a randomized, double-blind, placebo-controlled trial. *Drug Alcohol Depend*, *101*(1-2), 34-41. https://doi.org/10.1016/j.drugalcdep.2008.10.016

Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E., & Swarup, V. (2023). hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods*, *3*(6), 100498. https://doi.org/10.1016/j.crmeth.2023.100498

Moratalla, R., Vickers, E. A., Robertson, H. A., Cochran, B. H., & Graybiel, A. M. (1993). Coordinate expression of c-fos and jun B is induced in the rat striatum by cocaine. *J Neurosci*, *13*(2), 423-433. https://doi.org/10.1523/jneurosci.13-02-00423.1993

Morris, L. S., Kundu, P., Dowell, N., Mechelmans, D. J., Favre, P., Irvine, M. A., Robbins, T. W., Daw, N., Bullmore, E. T., Harrison, N. A., & Voon, V. (2016). Fronto-striatal organization: Defining functional and microstructural substrates of behavioural flexibility. *Cortex*, *74*, 118-133. https://doi.org/https://doi.org/10.1016/j.cortex.2015.11.004

Mustaquim, D., Jones, C. M., & Compton, W. M. (2021). Trends and correlates of cocaine use among adults in the United States, 2006-2019. *Addict Behav*, *120*, 106950. https://doi.org/10.1016/j.addbeh.2021.106950

Natarajaseenivasan, K., Cotto, B., Shanmughapriya, S., Lombardi, A. A., Datta, P. K., Madesh, M., Elrod, J. W., Khalili, K., & Langford, D. (2018). Astrocytic metabolic switch is a novel etiology for Cocaine and HIV-1 Tat-mediated neurotoxicity. *Cell Death & Disease*, *9*(4), 415. https://doi.org/10.1038/s41419-018-0422-3

Nestler, E. J. (2005). The neurobiology of cocaine addiction. *Sci Pract Perspect*, *3*(1), 4-10. https://doi.org/10.1151/spp05314

Nestler, E. J., Berhow, M. T., & Brodkin, E. S. (1996). Molecular mechanisms of drug addiction: adaptations in signal transduction pathways. *Mol Psychiatry*, *1*(3), 190-199.

Nestler, E. J., & Lüscher, C. (2019). The Molecular Basis of Drug Addiction: Linking Epigenetic to Synaptic and Circuit Mechanisms. *Neuron*, *102*(1), 48-59. https://doi.org/10.1016/j.neuron.2019.01.016

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., Hoang, C. D., Diehn, M., & Alizadeh, A. A. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*, *12*(5), 453-457. https://doi.org/10.1038/nmeth.3337

Nezhyva, M., Shahen-Zoabi, S., Kabirova, M., Bentov-Arava, E., Shalev, O., Andrén, P. E., Thornton, C., Yaka, R., Margulis, K., & Jansson, E. T. (2024). Spatial multiomic insights into acute cocaine exposure. *PNAS Nexus*, *3*(10), pgae458. https://doi.org/10.1093/pnasnexus/pgae458

Nichols, R. V., Rylaarsdam, L. E., O'Connell, B. L., Shipony, Z., Iremadze, N., Acharya, S. N., & Adey, A. C. (2025). Atlas-scale single-cell DNA methylation profiling with sciMETv3. *Cell Genomics*, *5*(1). https://doi.org/10.1016/j.xgen.2024.100726

NIDA. (2024). Drug Overdose Deaths: Facts and Figures *Retrieved from https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates on 2025, April 18*.

Niedzielska-Andres, E., Pomierny-Chamioło, L., Andres, M., Walczak, M., Knackstedt, L. A., Filip, M., & Przegaliński, E. (2021). Cocaine use disorder: A look at metabotropic glutamate receptors and glutamate transporters. *Pharmacol Ther*, *221*, 107797. https://doi.org/10.1016/j.pharmthera.2020.107797

Nielsen, D. A., Utrankar, A., Reyes, J. A., Simons, D. D., & Kosten, T. R. (2012). Epigenetics of drug abuse: predisposition or response. *Pharmacogenomics*, *13*(10), 1149-1160. https://doi.org/10.2217/pgs.12.94

Nocito Echevarria, M. A., Andrade Reis, T., Ruffo Capatti, G., Siciliano Soares, V., da Silveira, D. X., & Fidalgo, T. M. (2017). N-acetylcysteine for treating cocaine addiction - A systematic review. *Psychiatry Res*, *251*, 197-203. https://doi.org/10.1016/j.psychres.2017.02.024

Northcutt, A. L., Hutchinson, M. R., Wang, X., Baratta, M. V., Hiranita, T., Cochran, T. A., Pomrenze, M. B., Galer, E. L., Kopajtic, T. A., Li, C. M., Amat, J., Larson, G., Cooper, D. C., Huang, Y., O'Neill, C. E., Yin, H., Zahniser, N. R., Katz, J. L., Rice, K. C., . . . Watkins, L. R. (2015). DAT isn't all that: cocaine reward and reinforcement require Toll-like receptor 4 signaling. *Mol Psychiatry*, *20*(12), 1525-1537. https://doi.org/10.1038/mp.2014.177

Nuijten, M., Blanken, P., van de Wetering, B., Nuijen, B., van den Brink, W., & Hendriks, V. M. (2016). Sustained-release dexamfetamine in the treatment of chronic cocaine-dependent patients on heroin-assisted treatment: a randomised, double-blind, placebo-controlled trial. *Lancet*, *387*(10034), 2226-2234. https://doi.org/10.1016/s0140-6736(16)00205-1

Nuijten, M., Blanken, P., van den Brink, W., & Hendriks, V. (2014). Treatment of crack-cocaine dependence with topiramate: a randomized controlled

feasibility trial in The Netherlands. *Drug Alcohol Depend*, *138*, 177-184. https://doi.org/10.1016/j.drugalcdep.2014.02.024

Paliwal, P., Hyman, S. M., & Sinha, R. (2008). Craving predicts time to cocaine relapse: Further validation of the Now and Brief versions of the cocaine craving questionnaire. *Drug and Alcohol Dependence*, *93*(3), 252-259. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2007.10.002

Panov, A., Schonfeld, P., Dikalov, S., Hemendinger, R., Bonkovsky, H. L., & Brooks, B. R. (2009). The neuromediator glutamate, through specific substrate interactions, enhances mitochondrial ATP production and reactive oxygen species generation in nonsynaptic brain mitochondria. *J Biol Chem*, *284*(21), 14448-14456. https://doi.org/10.1074/jbc.M900985200

Pati, S., Angel, P., Drake, R. R., Wagner, J. J., & Cummings, B. S. (2019). Lipidomic changes in the rat hippocampus following cocaine conditioning, extinction, and reinstatement of drug-seeking. *Brain Behav*, *9*(12), e01451. https://doi.org/10.1002/brb3.1451

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*, *14*(4), 417-419. https://doi.org/10.1038/nmeth.4197

Pauli, W. M., O'Reilly, R. C., Yarkoni, T., & Wager, T. D. (2016). Regional specialization within the human striatum for diverse psychological functions. *Proceedings of the National Academy of Sciences*, *113*(7), 1907-1912. https://doi.org/doi:10.1073/pnas.1507610113

Peakman, M. C., Colby, C., Perrotti, L. I., Tekumalla, P., Carle, T., Ulery, P., Chao, J., Duman, C., Steffen, C., Monteggia, L., Allen, M. R., Stock, J. L., Duman, R. S., McNeish, J. D., Barrot, M., Self, D. W., Nestler, E. J., & Schaeffer, E. (2003). Inducible, brain region-specific expression of a dominant negative mutant of c-Jun in transgenic mice decreases sensitivity to cocaine. *Brain Res*, *970*(1-2), 73-86. https://doi.org/10.1016/s0006-8993(03)02230-3

Pelegí-Sisó, D., de Prado, P., Ronkainen, J., Bustamante, M., & González, J. R. (2021). methylclock: a Bioconductor package to estimate DNA methylation age. *Bioinformatics*, *37*(12), 1759-1760. https://doi.org/10.1093/bioinformatics/btaa825

Phillips, R. A., 3rd, Tuscher, J. J., Fitzgerald, N. D., Wan, E., Zipperly, M. E., Duke, C. G., Ianov, L., & Day, J. J. (2023). Distinct subpopulations of D1 medium spiny neurons exhibit unique transcriptional responsiveness to cocaine. *Mol Cell Neurosci*, *125*, 103849. https://doi.org/10.1016/j.mcn.2023.103849

Phipson, B., Maksimovic, J., & Oshlack, A. (2016). missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics*, *32*(2), 286-288. https://doi.org/10.1093/bioinformatics/btv560

Piazza, P. V., & Deroche-Gamonet, V. (2013). A multistep general theory of transition to addiction. *Psychopharmacology (Berl)*, *229*(3), 387-413. https://doi.org/10.1007/s00213-013-3224-4

Pierce, R. C., Fant, B., Swinford-Jackson, S. E., Heller, E. A., Berrettini, W. H., & Wimmer, M. E. (2018). Environmental, genetic and epigenetic contributions to cocaine addiction. *Neuropsychopharmacology*, *43*(7), 1471-1480. https://doi.org/10.1038/s41386-018-0008-x

Pignatelli, M., & Bonci, A. (2015). Role of Dopamine Neurons in Reward and Aversion: A Synaptic Plasticity Perspective. *Neuron*, *86*(5), 1145-1157. https://doi.org/10.1016/j.neuron.2015.04.015

Pohořalá, V., Enkel, T., Bartsch, D., Spanagel, R., & Bernardi, R. E. (2021). Sign- and goal-tracking score does not correlate with addiction-like behavior following prolonged cocaine self-administration. *Psychopharmacology (Berl)*, *238*(8), 2335-2346. https://doi.org/10.1007/s00213-021-05858-z

Poisel, E., Zillich, L., Streit, F., Frank, J., Friske, M. M., Foo, J. C., Mechawar, N., Turecki, G., Hansson, A. C., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2023). DNA methylation in cocaine use disorder–An epigenome-wide approach in the human prefrontal cortex [Original Research]. *Frontiers in Psychiatry*, *14*. https://doi.org/10.3389/fpsyt.2023.1075250

R Core Team. (2021). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna.* *https://www.R-project.org*.

R Core Team. (2022). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing. Vienna: R Foundation*.

Raffetti, E., Melas, P. A., Landgren, A. J., Andersson, F., Forsell, Y., Lavebratt, C., & Galanti, M. R. (2021). DNA methylation of the glucocorticoid receptor gene predicts substance use in adolescence: longitudinal data from over 1000 young individuals. *Transl Psychiatry*, *11*(1), 477. https://doi.org/10.1038/s41398-021-01601-6

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, *44*(W1), W160-W165. https://doi.org/10.1093/nar/gkw257

Ramos, C., Roberts, J. B., Jasso, K. R., Ten Eyck, T. W., Everett, T., Pozo, P., Setlow, B., & McIntyre, J. C. (2021). Neuron-specific cilia loss differentially alters locomotor responses to amphetamine in mice. *J Neurosci Res*, *99*(3), 827-842. https://doi.org/10.1002/jnr.24755

Rasooly, D., Giambartolomei, C., Peloso, G. M., Dashti, H., Ferolito, B. R., Golden, D., Horimoto, A., Pietzner, M., Farber-Eger, E. H., Wells, Q. S., Bini, G., Proietti, G., Tartaglia, G. G., Kosik, N. M., Wilson, P. W. F.,

Phillips, L. S., Munroe, P. B., Petersen, S. E., Cho, K., . . . Joseph, J. (2025). Large-scale multi-omics identifies drug targets for heart failure with reduced and preserved ejection fraction. *Nat Cardiovasc Res*, *4*(3), 293-311. https://doi.org/10.1038/s44161-025-00609-1

Reissner, K. J., Brown, R. M., Spencer, S., Tran, P. K., Thomas, C. A., & Kalivas, P. W. (2014). Chronic administration of the methylxanthine propentofylline impairs reinstatement to cocaine by a GLT-1-dependent mechanism. *Neuropsychopharmacology*, *39*(2), 499-506. https://doi.org/10.1038/npp.2013.223

Ribeiro, E. A., Scarpa, J. R., Garamszegi, S. P., Kasarskis, A., Mash, D. C., & Nestler, E. J. (2017). Gene Network Dysregulation in Dorsolateral Prefrontal Cortex Neurons of Humans with Cocaine Use Disorder. *Sci Rep*, *7*(1), 5412. https://doi.org/10.1038/s41598-017-05720-3

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Ritz, M. C., Lamb, R. J., Goldberg, S. R., & Kuhar, M. J. (1987). Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science*, *237*(4819), 1219-1223. https://doi.org/10.1126/science.2820058

Robinson, T. E., Gorny, G., Mitton, E., & Kolb, B. (2001). Cocaine self-administration alters the morphology of dendrites and dendritic spines in the nucleus accumbens and neocortex. *Synapse*, *39*(3), 257-266. https://doi.org/10.1002/1098-2396(20010301)39:3<257::Aid-syn1007>3.0.Co;2-1

Robison, A. J., & Nestler, E. J. (2011). Transcriptional and epigenetic mechanisms of addiction. *Nature Reviews Neuroscience*, *12*(11), 623-637. https://doi.org/10.1038/nrn3111

Rohart, F., Gautier, B., Singh, A., & Lê Cao, K.-A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Computational Biology*, *13*(11), e1005752. https://doi.org/10.1371/journal.pcbi.1005752

Rossi, A., Riccio, A., Coccia, M., Trotta, E., La Frazia, S., & Santoro, M. G. (2014). The Proteasome Inhibitor Bortezomib Is a Potent Inducer of Zinc Finger AN1-type Domain 2a Gene Expression: ROLE OF HEAT SHOCK FACTOR 1 (HSF1)-HEAT SHOCK FACTOR 2 (HSF2) HETEROCOMPLEXES*. *Journal of Biological Chemistry*, *289*(18), 12705-12715. https://doi.org/https://doi.org/10.1074/jbc.M113.513242

Roussel-Gervais, A., Naciri, I., Kirsh, O., Kasprzyk, L., Velasco, G., Grillo, G., Dubus, P., & Defossez, P.-A. (2017). Loss of the Methyl-CpG–Binding Protein ZBTB4 Alters Mitotic Checkpoint, Increases Aneuploidy, and

Promotes Tumorigenesis. *Cancer Research*, *77*(1), 62-73. https://doi.org/10.1158/0008-5472.CAN-16-1181

Russo, S. J., Wilkinson, M. B., Mazei-Robison, M. S., Dietz, D. M., Maze, I., Krishnan, V., Renthal, W., Graham, A., Birnbaum, S. G., Green, T. A., Robison, B., Lesselyong, A., Perrotti, L. I., Bolaños, C. A., Kumar, A., Clark, M. S., Neumaier, J. F., Neve, R. L., Bhakar, A. L., . . . Nestler, E. J. (2009). Nuclear Factor κB Signaling Regulates Neuronal Morphology and Cocaine Reward. *The Journal of Neuroscience*, *29*(11), 3529-3537. https://doi.org/10.1523/jneurosci.6173-08.2009

Sabe, M., Zhao, N., & Kaiser, S. (2021). A systematic review and meta-analysis of the prevalence of cocaine-induced psychosis in cocaine users. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *109*, 110263. https://doi.org/https://doi.org/10.1016/j.pnpbp.2021.110263

Safran, M., Rosen, N., Twik, M., BarShir, R., Stein, T. I., Dahary, D., Fishilevich, S., & Lancet, D. (2021). The GeneCards Suite. In I. Abugessaisa & T. Kasukawa (Eds.), *Practical Guide to Life Science Databases* (pp. 27-56). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-5812-9_2

Salery, M., Godino, A., & Nestler, E. J. (2021). Drug-activated cells: From immediate early genes to neuronal ensembles in addiction. *Adv Pharmacol*, *90*, 173-216. https://doi.org/10.1016/bs.apha.2020.09.006

Saunders, J. B., Conigrave, K. M., Latt, N. C., Nutt, D. J., Marshall, E. J., Ling, W., & Higuchi, S. (2016). *Addiction Medicine*. Oxford University Press. https://doi.org/10.1093/med/9780198714750.001.0001

Savell, K. E., Tuscher, J. J., Zipperly, M. E., Duke, C. G., Phillips, R. A., 3rd, Bauman, A. J., Thukral, S., Sultan, F. A., Goska, N. A., Ianov, L., & Day, J. J. (2020). A dopamine-induced gene expression signature regulates neuronal function and cocaine response. *Sci Adv*, *6*(26), eaba4221. https://doi.org/10.1126/sciadv.aba4221

Schmidt, H. D., & Pierce, R. C. (2010). Cocaine-induced neuroadaptations in glutamate transmission: potential therapeutic targets for craving and addiction. *Ann N Y Acad Sci*, *1187*, 35-75. https://doi.org/10.1111/j.1749-6632.2009.05144.x

Schote, A. B., Jäger, K., Kroll, S. L., Vonmoos, M., Hulka, L. M., Preller, K. H., Meyer, J., Grünblatt, E., & Quednow, B. B. (2019). Glucocorticoid receptor gene variants and lower expression of NR3C1 are associated with cocaine use. *Addict Biol*, *24*(4), 730-742. https://doi.org/10.1111/adb.12632

Schwartz, E. K. C., Wolkowicz, N. R., De Aquino, J. P., MacLean, R. R., & Sofuoglu, M. (2022). Cocaine Use Disorder (CUD): Current Clinical Perspectives. *Subst Abuse Rehabil*, *13*, 25-46. https://doi.org/10.2147/sar.S337338

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, *13*(11), 2498-2504. https://doi.org/10.1101/gr.1239303

Shen L, S. I. (2023). GeneOverlap: Test and visualize gene overlaps. R package version 1.36.0.

Shippy, D. C., & Ulland, T. K. (2022). Exploring the zinc-related transcriptional landscape in Alzheimer's disease. *IBRO Neurosci Rep*, *13*, 31-37. https://doi.org/10.1016/j.ibneur.2022.06.002

Shu, C., Jaffe, A. E., Sabunciyan, S., Ji, H., Astemborski, J., Sun, J., Bakulski, K. M., Sosnowski, D. W., Mehta, S. H., Kirk, G. D., & Maher, B. S. (2022). Epigenome-wide association analyses of active injection drug use. *Drug Alcohol Depend*, *235*, 109431. https://doi.org/10.1016/j.drugalcdep.2022.109431

Shu, C., Sosnowski, D. W., Tao, R., Deep-Soboslay, A., Kleinman, J. E., Hyde, T. M., Jaffe, A. E., Sabunciyan, S., & Maher, B. S. (2021). Epigenome-wide study of brain DNA methylation following acute opioid intoxication. *Drug Alcohol Depend*, *221*, 108658. https://doi.org/10.1016/j.drugalcdep.2021.108658

Sil, S., Niu, F., Tom, E., Liao, K., Periyasamy, P., & Buch, S. (2019). Cocaine Mediated Neuroinflammation: Role of Dysregulated Autophagy in Pericytes. *Mol Neurobiol*, *56*(5), 3576-3590. https://doi.org/10.1007/s12035-018-1325-0

Singh, A., Shannon, C. P., Gautier, B., Rohart, F., Vacher, M., Tebbutt, S. J., & Lê Cao, K.-A. (2019). DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*, *35*(17), 3055-3062. https://doi.org/10.1093/bioinformatics/bty1054

Smart, R. G. (1991). Crack Cocaine Use: A Review of Prevalence and Adverse Effects. *The American Journal of Drug and Alcohol Abuse*, *17*(1), 13-26. https://doi.org/10.3109/00952999108992806

Smith, J. A., Qiu, M., Guo, H., Kunko, P. M., & Robinson, S. E. (1995). Cocaine increases extraneuronal levels of aspartate and glutamate in the nucleus accumbens. *Brain Research*, *683*(2), 264-269. https://doi.org/https://doi.org/10.1016/0006-8993(95)00383-2

Soneson, C., Love, M. I., & Robinson, M. D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*, *4*, 1521. https://doi.org/10.12688/f1000research.7563.2

Spanagel, R., Bach, P., Banaschewski, T., Beck, A., Bermpohl, F., Bernardi, R. E., Beste, C., Deserno, L., Durstewitz, D., Ebner-Priemer, U., Endrass, T., Ersche, K. D., Feld, G., Gerchen, M. F., Gerlach, B., Goschke, T., Hansson, A. C., Heim, C., Kiebel, S., . . . Heinz, A. (2024). The ReCoDe addiction

research consortium: Losing and regaining control over drug intake-Findings and future perspectives. *Addict Biol*, *29*(7), e13419. https://doi.org/10.1111/adb.13419

Stewart, A. F., Fulton, S. L., & Maze, I. (2021). Epigenetics of Drug Addiction. *Cold Spring Harb Perspect Med*, *11*(7). https://doi.org/10.1101/cshperspect.a040253

Stiltner, B., Pietrzak, R. H., Tylee, D. S., Nunez, Y. Z., Adhikari, K., Kranzler, H. R., Gelernter, J., & Polimanti, R. (2023). Polysubstance addiction patterns among 7,989 individuals with cocaine use disorder. *iScience*, *26*(8), 107336. https://doi.org/https://doi.org/10.1016/j.isci.2023.107336

Su, Y., Chen, D., Yuan, D., Lausted, C., Choi, J., Dai, C. L., Voillet, V., Duvvuri, V. R., Scherler, K., Troisch, P., Baloni, P., Qin, G., Smith, B., Kornilov, S. A., Rostomily, C., Xu, A., Li, J., Dong, S., Rothchild, A., . . . Heath, J. R. (2020). Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. *Cell*, *183*(6), 1479-1495.e1420. https://doi.org/https://doi.org/10.1016/j.cell.2020.10.037

Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Gould, J., Davis, J. F., Tubelli, A. A., Asiedu, J. K., Lahr, D. L., Hirschman, J. E., Liu, Z., Donahue, M., Julian, B., Khan, M., Wadden, D., Smith, I. C., Lam, D., . . . Golub, T. R. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*, *171*(6), 1437-1452.e1417. https://doi.org/10.1016/j.cell.2017.10.049

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, *102*(43), 15545-15550. https://doi.org/10.1073/pnas.0506580102

Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics Data Integration, Interpretation, and Its Application. *Bioinform Biol Insights*, *14*, 1177932219899051. https://doi.org/10.1177/1177932219899051

Suderman, M., Staley, J. R., French, R., Arathimos, R., Simpkin, A., & Tilling, K. (2018). dmrff: identifying differentially methylated regions efficiently with power and control. *bioRxiv*, 508556. https://doi.org/10.1101/508556

Sun, C., Nold, A., Fusco, C. M., Rangaraju, V., Tchumatchenko, T., Heilemann, M., & Schuman, E. M. (2021). The prevalence and specificity of local protein synthesis during neuronal synaptic plasticity. *Science Advances*, *7*(38), eabj0790. https://doi.org/doi:10.1126/sciadv.abj0790

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any

sequenced genome of interest. *Nucleic Acids Res*, *51*(D1), D638-d646. https://doi.org/10.1093/nar/gkac1000

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377-382. https://doi.org/10.1038/nmeth.1315

Teague, C. D., & Nestler, E. J. (2022). Key transcription factors mediating cocaine-induced plasticity in the nucleus accumbens. *Molecular Psychiatry*, *27*(1), 687-709. https://doi.org/10.1038/s41380-021-01163-5

Thanos, P. K., Michaelides, M., Benveniste, H., Wang, G. J., & Volkow, N. D. (2008). The effects of cocaine on regional brain glucose metabolism is attenuated in dopamine transporter knockout mice. *Synapse*, *62*(5), 319-324. https://doi.org/10.1002/syn.20503

Thiriet, N., Aunis, D., & Zwiller, J. (2000). C-fos and egr-1 immediate-early gene induction by cocaine and cocaethylene in rat brain: a comparative study. *Ann N Y Acad Sci*, *914*, 46-57. https://doi.org/10.1111/j.1749-6632.2000.tb05182.x

Thompson, A., Wölmer, N., Koncarevic, S., Selzer, S., Böhm, G., Legner, H., Schmid, P., Kienle, S., Penning, P., Höhle, C., Berfelde, A., Martinez-Pinna, R., Farztdinov, V., Jung, S., Kuhn, K., & Pike, I. (2019). TMTpro: Design, Synthesis, and Initial Evaluation of a Proline-Based Isobaric 16-Plex Tandem Mass Tag Reagent Set. *Anal Chem*, *91*(24), 15941-15950. https://doi.org/10.1021/acs.analchem.9b04474

Thornton, C., Grad, E., & Yaka, R. (2021). The role of mitochondria in cocaine addiction. *Biochemical Journal*, *478*(4), 749-764. https://doi.org/10.1042/BCJ20200615

Tian, W., Wang, J., Zhang, K., Teng, H., Li, C., Szyf, M., Sun, Z. S., & Zhao, M. (2016). Demethylation of c-MYB binding site mediates upregulation of Bdnf IV in cocaine-conditioned place preference. *Scientific Reports*, *6*(1), 22087. https://doi.org/10.1038/srep22087

Tian, W., Zhao, M., Li, M., Song, T., Zhang, M., Quan, L., Li, S., & Sun, Z. S. (2012). Reversal of cocaine-conditioned place preference through methyl supplementation in mice: altering global DNA methylation in the prefrontal cortex. *PLoS One*, *7*(3), e33435. https://doi.org/10.1371/journal.pone.0033435

Tondo, L. P., Viola, T. W., Fries, G. R., Kluwe-Schiavon, B., Rothmann, L. M., Cupertino, R., Ferreira, P., Franco, A. R., Lane, S. D., Stertz, L., Zhao, Z., Hu, R., Meyer, T., Schmitz, J. M., Walss-Bass, C., & Grassi-Oliveira, R. (2021). White matter deficits in cocaine use disorder: convergent evidence from in vivo diffusion tensor imaging and ex vivo proteomic analysis. *Transl Psychiatry*, *11*(1), 252. https://doi.org/10.1038/s41398-021-01367-x

Tran, M. N., Maynard, K. R., Spangler, A., Huuki, L. A., Montgomery, K. D., Sadashivaiah, V., Tippani, M., Barry, B. K., Hancock, D. B., Hicks, S. C., Kleinman, J. E., Hyde, T. M., Collado-Torres, L., Jaffe, A. E., & Martinowich, K. (2021). Single-nucleus transcriptome analysis reveals cell-type-specific molecular signatures across reward circuitry in the human brain. *Neuron*, *109*(19), 3088-3103.e3085. https://doi.org/https://doi.org/10.1016/j.neuron.2021.09.001

Trotter, J. H., Dargaei, Z., Sclip, A., Essayan-Perez, S., Liakath-Ali, K., Raju, K., Nabet, A., Liu, X., Wöhr, M., & Südhof, T. C. (2021). Compartment-Specific Neurexin Nanodomains Orchestrate Tripartite Synapse Assembly. *bioRxiv*, 2020.2008.2021.262097. https://doi.org/10.1101/2020.08.21.262097

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, *1*(1), 59. https://doi.org/10.1038/s43586-021-00056-9

UNODC. (2022). World Drug Report 2022, New York, NY: United Nations publication.

UNODC. (2023a). Global report on cocaine : 2023 local dynamics, global challenges.

UNODC. (2023b). World Drug Report 2023 (United Nations Publication).

Vaillancourt, K., Chen, G. G., Fiori, L., Maussion, G., Yerko, V., Théroux, J. F., Ernst, C., Labonté, B., Calipari, E., Nestler, E. J., Nagy, C., Mechawar, N., Mash, D. C., & Turecki, G. (2021). Methylation of the tyrosine hydroxylase gene is dysregulated by cocaine dependence in the human striatum. *iScience*, *24*(10), 103169. https://doi.org/10.1016/j.isci.2021.103169

Vaillancourt, K., Ernst, C., Mash, D., & Turecki, G. (2017). DNA Methylation Dynamics and Cocaine in the Brain: Progress and Prospects. *Genes (Basel)*, *8*(5). https://doi.org/10.3390/genes8050138

Vaillancourt, K., Yang, J., Chen, G. G., Yerko, V., Théroux, J. F., Aouabed, Z., Lopez, A., Thibeault, K. C., Calipari, E. S., Labonté, B., Mechawar, N., Ernst, C., Nagy, C., Forné, T., Nestler, E. J., Mash, D. C., & Turecki, G. (2021). Cocaine-related DNA methylation in caudate neurons alters 3D chromatin structure of the IRXA gene cluster. *Mol Psychiatry*, *26*(7), 3134-3151. https://doi.org/10.1038/s41380-020-00909-x

Van Booven, D., Mengying, L., Sunil Rao, J., Blokhin, I. O., Dayne Mayfield, R., Barbier, E., Heilig, M., & Wahlestedt, C. (2021). Alcohol use disorder causes global changes in splicing in the human brain. *Transl Psychiatry*, *11*(1), 2. https://doi.org/10.1038/s41398-020-01163-z

Velten, B., Braunger, J. M., Argelaguet, R., Arnol, D., Wirbel, J., Bredikhin, D., Zeller, G., & Stegle, O. (2022). Identifying temporal and spatial patterns of

variation from multimodal data using MEFISTO. *Nat Methods*, *19*(2), 179-186. https://doi.org/10.1038/s41592-021-01343-9

Verhulst, B., Neale, M. C., & Kendler, K. S. (2015). The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol Med*, *45*(5), 1061-1072. https://doi.org/10.1017/s0033291714002165

Viola, T. W., Heberle, B. A., Zaparte, A., Sanvicente-Vieira, B., Wainer, L. M., Fries, G. R., Walss-Bass, C., & Grassi-Oliveira, R. (2019). Peripheral blood microRNA levels in females with cocaine use disorder. *J Psychiatr Res*, *114*, 48-54. https://doi.org/10.1016/j.jpsychires.2019.03.028

Volkow, N. D., Fowler, J. S., & Wang, G. J. (2003). The addicted human brain: insights from imaging studies. *J Clin Invest*, *111*(10), 1444-1451. https://doi.org/10.1172/jci18533

Volkow, N. D., Koob, G. F., & McLellan, A. T. (2016). Neurobiologic Advances from the Brain Disease Model of Addiction. *New England Journal of Medicine*, *374*(4), 363-371. https://doi.org/doi:10.1056/NEJMra1511480

Volkow, N. D., Michaelides, M., & Baler, R. (2019). The Neuroscience of Drug Reward and Addiction. *Physiological Reviews*, *99*(4), 2115-2140. https://doi.org/10.1152/physrev.00014.2018

Vollstädt-Klein, S., Wichert, S., Rabinstein, J., Bühler, M., Klein, O., Ende, G., Hermann, D., & Mann, K. (2010). Initial, habitual and compulsive alcohol use is characterized by a shift of cue processing from ventral to dorsal striatum. *Addiction*, *105*(10), 1741-1749. https://doi.org/10.1111/j.1360-0443.2010.03022.x

Vornholt, E., Liharska, L. E., Cheng, E., Hashemi, A., Park, Y. J., Ziafat, K., Wilkins, L., Silk, H., Linares, L. M., Thompson, R. C., Sullivan, B., Moya, E., Nadkarni, G. N., Sebra, R., Schadt, E. E., Kopell, B. H., Charney, A. W., & Beckmann, N. D. (2024). Characterizing cell type specific transcriptional differences between the living and postmortem human brain. *medRxiv*. https://doi.org/10.1101/2024.05.01.24306590

Wagner, F. A., & Anthony, J. C. (2002). From first drug use to drug dependence; developmental periods of risk for dependence upon marijuana, cocaine, and alcohol. *Neuropsychopharmacology*, *26*(4), 479-488. https://doi.org/10.1016/s0893-133x(01)00367-0

Walker, C. D., Risher, W. C., & Risher, M. L. (2020). Regulation of Synaptic Development by Astrocyte Signaling Factors and Their Emerging Roles in Substance Abuse. *Cells*, *9*(2). https://doi.org/10.3390/cells9020297

Wang, J., Holt, L. M., Huang, H. H., Sesack, S. R., Nestler, E. J., & Dong, Y. (2022). Astrocytes in cocaine addiction and beyond. *Mol Psychiatry*, *27*(1), 652-668. https://doi.org/10.1038/s41380-021-01080-7

Weiss, F., Markou, A., Lorang, M. T., & Koob, G. F. (1992). Basal extracellular dopamine levels in the nucleus accumbens are decreased during cocaine withdrawal after unlimited-access self-administration. *Brain Res*, *593*(2), 314-318. https://doi.org/10.1016/0006-8993(92)91327-b

Werner, T., Sweetman, G., Savitski, M. F., Mathieson, T., Bantscheff, M., & Savitski, M. M. (2014). Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal Chem*, *86*(7), 3594-3601. https://doi.org/10.1021/ac500140s

World Health Organization. (2019). International Statistical Classification of Diseases and Related
Health Problems (11th ed.). https://icd.who.int/

Wright, K. N., Hollis, F., Duclot, F., Dossat, A. M., Strong, C. E., Francis, T. C., Mercer, R., Feng, J., Dietz, D. M., Lobo, M. K., Nestler, E. J., & Kabbaj, M. (2015). Methyl supplementation attenuates cocaine-seeking behaviors and cocaine-induced c-Fos activation in a DNA methylation-dependent manner. *J Neurosci*, *35*(23), 8948-8958. https://doi.org/10.1523/jneurosci.5227-14.2015

Xu, S.-J., Lombroso, S. I., Fischer, D. K., Carpenter, M. D., Marchione, D. M., Hamilton, P. J., Lim, C. J., Neve, R. L., Garcia, B. A., Wimmer, M. E., Pierce, R. C., & Heller, E. A. (2021). Chromatin-mediated alternative splicing regulates cocaine-reward behavior. *Neuron*, *109*(18), 2943-2966.e2948. https://doi.org/10.1016/j.neuron.2021.08.008

Yager, L. M., Garcia, A. F., Wunsch, A. M., & Ferguson, S. M. (2015). The ins and outs of the striatum: role in drug addiction. *Neuroscience*, *301*, 529-541. https://doi.org/10.1016/j.neuroscience.2015.06.033

Yamazaki, Y., Zhao, N., Caulfield, T. R., Liu, C.-C., & Bu, G. (2019). Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. *Nature Reviews Neurology*, *15*(9), 501-518. https://doi.org/10.1038/s41582-019-0228-7

Yang, K. C., & Gorski, S. M. (2022). Protocol for analysis of RNA-sequencing and proteome profiling data for subgroup identification and comparison. *STAR Protocols*, *3*(2), 101283. https://doi.org/https://doi.org/10.1016/j.xpro.2022.101283

Yao, W., Huang, J., & He, H. (2019). Over-expressed LOC101927196 suppressed oxidative stress levels and neuron cell proliferation in a rat model of autism through disrupting the Wnt signaling pathway by targeting FZD3. *Cell Signal*, *62*, 109328. https://doi.org/10.1016/j.cellsig.2019.05.013

Young, M. D., & Behjati, S. (2020). SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *GigaScience*, *9*(12), giaa151. https://doi.org/10.1093/gigascience/giaa151

Yu. (2024). enrichplot: Visualization of Functional Enrichment Result. R package version 1.24.4, https://yulab-smu.top/biomedical-knowledge-mining-book/.

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics*, *16*(5), 284-287. https://doi.org/10.1089/omi.2011.0118

Yu, Q., & He, Z. (2017). Comprehensive investigation of temporal and autism-associated cell type composition-dependent and independent gene expression changes in human brains. *Scientific Reports*, *7*(1), 4121. https://doi.org/10.1038/s41598-017-04356-7

Yugi, K., Kubota, H., Hatano, A., & Kuroda, S. (2016). Trans-Omics: How To Reconstruct Biochemical Networks Across Multiple ‘Omic’ Layers. *Trends in Biotechnology*, *34*(4), 276-290. https://doi.org/10.1016/j.tibtech.2015.12.013

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., . . . Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*(1), 14049. https://doi.org/10.1038/ncomms14049

Zhou, J. L., de Guglielmo, G., Ho, A. J., Kallupi, M., Pokhrel, N., Li, H.-R., Chitre, A. S., Munro, D., Mohammadi, P., Carrette, L. L. G., George, O., Palmer, A. A., McVicker, G., & Telese, F. (2023). Single-nucleus genomics in outbred rats with divergent cocaine addiction-like behaviors reveals changes in amygdala GABAergic inhibition. *Nature Neuroscience*, *26*(11), 1868-1879. https://doi.org/10.1038/s41593-023-01452-y

Zhou, Z., Yuan, Q., Mash, D. C., & Goldman, D. (2011). Substance-specific and shared transcription and epigenetic changes in the human hippocampus chronically exposed to cocaine and alcohol. *Proceedings of the National Academy of Sciences*, *108*(16), 6626-6631. https://doi.org/10.1073/pnas.1018514108

Zillich, E., Belschner, H., Avetyan, D., Andrade-Brito, D., Martínez-Magaña, J. J., Frank, J., Mechawar, N., Turecki, G., Cabana-Domínguez, J., Fernàndez-Castillo, N., Cormand, B., Montalvo-Ortiz, J. L., Nöthen, M. M., Hansson, A. C., Rietschel, M., Spanagel, R., Witt, S. H., & Zillich, L. (2024). Multi-omics profiling of DNA methylation and gene expression alterations in human cocaine use disorder. *Transl Psychiatry*, *14*(1), 428. https://doi.org/10.1038/s41398-024-03139-9

Zillich, L., Artioli, A., Pohořalá, V., Zillich, E., Stertz, L., Belschner, H., Jabali, A., Frank, J., Streit, F., Avetyan, D., Völker, M. P., Müller, S., Hansson, A. C., Meyer, T. D., Rietschel, M., Ladewig, J., Spanagel, R., Oliveira, A. M. M., Walss-Bass, C., . . . Witt, S. H. (2025). Cell type-specific multi-omics

analysis of cocaine use disorder in the human caudate nucleus. *Nature Communications*, *16*(1), 3381. https://doi.org/10.1038/s41467-025-57339-y

Zillich, L., Frank, J., Streit, F., Friske, M. M., Foo, J. C., Sirignano, L., Heilmann-Heimbach, S., Dukal, H., Degenhardt, F., Hoffmann, P., Hansson, A. C., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2022). Epigenome-wide association study of alcohol use disorder in five brain regions. *Neuropsychopharmacology*, *47*(4), 832-839. https://doi.org/10.1038/s41386-021-01228-7

Zillich, L., Poisel, E., Frank, J., Foo, J. C., Friske, M. M., Streit, F., Sirignano, L., Heilmann-Heimbach, S., Heimbach, A., Hoffmann, P., Degenhardt, F., Hansson, A. C., Bakalkin, G., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2022). Multi-omics signatures of alcohol use disorder in the dorsal and ventral striatum. *Translational Psychiatry*, *12*(1), 190. https://doi.org/10.1038/s41398-022-01959-1

8 RELATED PUBLICATIONS

Zillich, L., Cetin, M., Hummel, E. M., Poisel, E., Fries, G. R., Frank, J., Streit, F., Foo, J. C., Sirignano, L., Friske, M. M., Lenz, B., Hoffmann, S., Adorjan, K., Kiefer, F., Bakalkin, G., Hansson, A. C., Lohoff, F. W., Kärkkäinen, O., Kok, E., . . . Witt, S. H. (2024). Biological aging markers in blood and brain tissue indicate age acceleration in alcohol use disorder. *Alcohol Clin Exp Res (Hoboken)*, *48*(2), 250-259. https://doi.org/10.1111/acer.15241

Zillich, L., Poisel, E., Frank, J., Foo, J. C., Friske, M. M., Streit, F., Sirignano, L., Heilmann-Heimbach, S., Heimbach, A., Hoffmann, P., Degenhardt, F., Hansson, A. C., Bakalkin, G., Nöthen, M. M., Rietschel, M., Spanagel, R., & Witt, S. H. (2022). Multi-omics signatures of alcohol use disorder in the dorsal and ventral striatum. *Translational Psychiatry*, *12*(1), 190. https://doi.org/10.1038/s41398-022-01959-1

Zillich, L., Poisel, E., Streit, F., Frank, J., Fries, G. R., Foo, J. C., Friske, M. M., Sirignano, L., Hansson, A. C., Nöthen, M. M., Witt, S. H., Walss-Bass, C., Spanagel, R., & Rietschel, M. (2022). Epigenetic Signatures of Smoking in Five Brain Regions. *J Pers Med*, *12*(4). https://doi.org/10.3390/jpm12040566

Zillich, L., Artioli, A., Pohořalá, V., Zillich, E., Stertz, L., Belschner, H., Jabali, A., Frank, J., Streit, F., Avetyan, D., Völker, M. P., Müller, S., Hansson, A. C., Meyer, T. D., Rietschel, M., Ladewig, J., Spanagel, R., Oliveira, A. M. M., Walss-Bass, C., . . . Witt, S. H. (2025). Cell type-specific multi-omics analysis of cocaine use disorder in the human caudate nucleus. *Nature Communications*, *16*(1), 3381. https://doi.org/10.1038/s41467-025-57339-y

9 CURRICULUM VITAE

PERSONALIEN

Name und Vorname:	Eric Marcel Zillich (geb. Poisel)
Geburtsdatum:	13.08.1999
Geburtsort:	Crailsheim

SCHULISCHER WERDEGANG

2009 – 2016	Albert-Schweitzer-Gymnasium Crailsheim
30.06.2016	Abitur

UNIVERSITÄRER WERDEGANG

WS2016/2017	Beginn des Bachelorstudiums der Biochemie An der Ruprecht-Karls-Universität Heidelberg
SS2019	Bachelorarbeit: “Effects of Glial Co-Culture on a Murine Astrocytoma Model”
02.07.2019	Bachelor of Science (B.Sc.), Note 1.1
2019 – 2021	Masterstudium der Biochemie An der Ruprecht-Karls-Universität Heidelberg
SS2021- WS2021/2022	Masterarbeit: „DNA Methylation Signatures of Smoking in the Brain“
10.12.2021	Master of Science (M.Sc.), Note 1.0
SS2022	Beginn des Studiums der Humanmedizin An der Johannes-Gutenberg-Universität Mainz
03.04.2025	1. Abschnitt der Ärztlichen Prüfung, Note 1.0

10 DANKSAGUNG

Zuerst möchte ich mich bei meiner Doktormutter PD Dr. Stephanie Witt für die Überlassung des Themas, die konstruktiven Gespräche, und das Feedback über alle Aspekte der Promotion hinweg bedanken. Vielen Dank auch für die unterstützende Arbeitsatmosphäre, in der ich stets darin bestärkt wurde, neue Ideen und Methoden für die Studien zu evaluieren und anzuwenden.

Bei Prof. Dr. Marcella Rietschel möchte ich mich für die kontinuierliche Unterstützung danken, die ich schon seit dem Eintreten in die Arbeitsgruppe, damals noch im Rahmen eines Forschungspraktikums, erfahren durfte. Neben den zahlreichen Ideen und den stets konstruktiven kritischen Einwänden bei der Diskussion von Methodik und Ergebnissen, sind mir auch die Unterhaltungen zu philosophischen Aspekten der Psychiatrie noch in guter Erinnerung.

Ein großer Dank gilt auch Prof. Dr. Rainer Spanagel für die fachliche Unterstützung, den Antrieb, die eigene Forschung stets auf die nächste Ebene zu bringen, sowie für die hilfreichen Überlegungen und Ansätze, die sich während der TACs, aber auch in anderen Gesprächen ergeben haben.

Bei PD Dr. Christian Witt bedanke ich mich für die zahlreichen Gespräche fachlicher Natur, insbesondere im Bereich Proteomics, aber auch in anderen Themenbereichen, die zumeist recht unterhaltsam waren.

Den aktuellen und ehemaligen Kolleginnen und Kollegen der Abteilung für Genetische Epidemiologie in der Psychiatrie am ZI – Dr. Diana Avetyan, Hanna Belschner, Daniela Distel, Helene Dukal, Mariana Espinola, Dr. Jerome Foo, Dr. Josef Frank, Dr. Svenja Müller, Dr. Thao Nguyen, Dr. Lea Sirignano, Dr. Fabian Streit und Maja Völker – danke ich für ihre kontinuierliche Unterstützung in den letzten Jahren, für gute Gespräche, und für ihr wertvolles Feedback in Progress Reports und Journal Clubs.

Vielen Dank an die Ko-Autorinnen und Ko-Autoren der drei Studien, die sowohl innerhalb des ZI, beispielsweise im Institut für Psychopharmakologie und dem Hector Institut für Translationale Hirnforschung, als auch in nationalen und internationalen Kollaborationen maßgeblich zum Erfolg des Projekts beigetragen haben.

Ein Dank auch an das Team der Biobank des ZI, insbesondere an Madeleine Hornberger, für das zuverlässige Probenhandling.

Dem Team der Forschungs-IT des ZI möchte ich für ihren schnellen und hilfreichen Support bei Fragen und für ihre zuverlässige Lösung bei IT-Problemen danken.

Vielen Dank an metis und gaia für die unermüdliche Arbeit, auch nachts und an Wochenenden.

Ein besonders großer Dank gilt den Gewebespendern, die ihren Körper nach dem Tod der Wissenschaft zur Verfügung gestellt haben. Dies erachte ich als einen großen Dienst an der Menschheit, der maßgeblich dazu beiträgt, durch Fortschritte in der Forschung zukünftigen Generationen ein besseres Leben zu ermöglichen.

Meiner Familie und Freunden möchte ich für ihre langjährige und teils schon lebenslange Unterstützung danken. Insbesondere meinen Eltern und Großeltern danke ich für den Freiraum, seit jeher meiner Neugier und meinem Interesse an der Natur nachgehen zu dürfen und dadurch die ein oder anderen Spezialinteressen entwickeln zu dürfen, von denen glücklicherweise nicht mehr alle bis heute erhalten blieben.

Lea, für alles.