DNA METHYLATION AND COCAINE DEPENDENCE IN THE HUMAN STRIATUM

Kathryn Lynn Vaillancourt

Integrated Program in Neuroscience

Department of Neurology and Neurosurgery

McGill University, Montreal

December 2020

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

"They can't take an education away from you, darling — you'll always have that with you."

- Marjorie Kathryn Fadden

TABLE OF CONTENTS

ABSTRACT

Cocaine use disorder (CUD) is a chronic relapsing brain disease that involves cycles of intoxication, withdrawal, and preoccupation. CUD impacts 1-2% of the population worldwide, and although there are multiple lines of evidence that report lasting changes in neurobiology, the mechanisms that accompany the development and maintenance of the disorder are not well understood. Work in animal models suggests that epigenetic mechanisms, including histone post translational modifications and DNA methylation, might be important regulators of cocaine-related changes to cell functioning, but to date, little is known about the relationship between cocaine dependence and epigenetics in humans.

This thesis presents the first methylome-wide investigations in human post-mortem brain, using multiple cohorts of banked samples, from two addiction-relevant brain regions; the nucleus accumbens which is highly implicated in the motivational effects of drug use, and the caudate nucleus which is involved in compulsive drug-seeking behaviors. We use reduced representation bisulfite sequencing (RRBS) and identified and characterized over 100 differentially methylated regions (DMRs) in both brain areas. We also report on two novel intragenic regulatory elements which are differentially methylated in the neurons of cocaine dependent subjects — a CTCF binding element within the Iroquois homeobox 2 gene (*IRX2*) which regulates local three-dimensional chromatin structure, and an EGR1 binding element within the tyrosine hydroxylase gene (*TH*) which acts as a methylation-sensitive enhancer of gene expression. This work represents an important first step towards understanding the role of DNA methylation in the pathophysiology of CUD in humans and provides numerous avenues for further investigation.

RESUMÉ

Le trouble lié à l'usage de la cocaïne (CUD) est une maladie cérébrale chronique récurrente qui implique des cycles d'intoxication, de sevrage et de préoccupation. La CUD affecte 1 à 2% de la population dans le monde, et bien qu'il existe de multiples sources de preuves rapportant des changements durables en neurobiologie, les mécanismes qui accompagnent le développement et le maintien de la maladie ne sont pas bien compris. L'études sur des modèles animaux suggèrent que les mécanismes épigénétiques, y compris les modifications post-traductionnelles des histones et la méthylation de l'ADN, pourraient être des régulateurs importants des changements liés à la cocaïne dans le fonctionnement des cellules, mais à ce jour, on en sait peu sur la relation entre la dépendance à la cocaïne et l'épigénétique chez les humains.

Cette thèse présente les premières enquêtes à l'échelle du méthylome dans le cerveau post-mortem humain, en utilisant plusieurs cohortes d'échantillons en banque, provenant de deux régions du cerveau liées à la dépendance; le noyau accumbens qui est fortement impliqué dans les effets motivationnels de la consommation de drogue, et le noyau caudé qui est impliqué dans les comportements compulsifs de recherche de drogue. Nous utilisons la méthode «reduced representation bisulfite sequencing» (RRBS) et nous avons identifié et caractérisé plus de 100 régions méthylées différentiellement (RMD) dans les deux zones du cerveau. Nous rapportons également sur deux nouveaux éléments régulateurs intragéniques qui sont méthylés de manière différentielle dans les neurones des sujets dépendants de la cocaïne - un élément de liaison CTCF dans le gène Iroquois homeobox 2 (*IRX2*) qui régule la structure de la chromatine tridimensionnelle locale, et un élément de liaison EGR1 dans le gène de tyrosine hydroxylase (*TH*) qui agit comme un amplificateur sensible à la méthylation de l'expression génique. Ce travail représente une première étape importante vers la compréhension du rôle de la méthylation de l'ADN dans la physiopathologie de la CUD chez l'humain et offre de nombreuses pistes pour une enquête plus approfondie.

ACKNOWLEDGEMENTS

Although this document bears my name, the work herein would not exist without the contribution of many, many others. To my thesis supervisor, Gustavo Turecki, thank you for agreeing to support my training and guiding me through these last eight years of discovery. To the members of my advisory committee, Cecilia Flores and Alain Gratton, thank you for your continued insight and feedback. Thank you to Jennie Yang, Volodymyr Yerko and Gary G Chen for the innumerable skills you've taught me, both in the lab and in life. I will always be grateful to have had the opportunity to work alongside the brilliant group of scientists at the McGill Group for Suicide Studies. Particular thanks to Corina Nagy, Cristiana Cruceanu and Marissa Maheu for giving me exceptionally strong examples of what it means to be a woman in science. To Lea Perret, Daniel Almeida, Jeffrey Gross, and Juan Pablo Lopez – thank you for believing in my potential, even when I lost faith in myself, and for being such enthusiastic supporters of all my wild ideas.

I also want to acknowledge the massive contribution that my family has made to my success. To my grandparents Marj, Ray, Corrie, and Roy, thank you for encouraging me to explore my passions and giving me the space to be creative. To my parents Brenda and François, there aren't enough words to thank you for the emotional and financial support you've provided while I've been on this path. Dad, I will always count on you to help me understand my value and to find the positives in even the darkest situation. Mom, I would literally not have survived this without you, and I owe my strength and resilience to your example and guidance.

Finally, thank you to Scarlett and Julianne for your unbridled imagination, your incredibly thoughtful questions, and your unstoppable sense of wonder that reminds me of all the reasons why I started this journey in the first place.

PREFACE TO THE THESIS

This thesis is presented in the manuscript-based format for Doctoral Theses, as described in the Thesis Preparation Guidelines by the Department of Graduate and Postdoctoral Studies at McGill University. The work described here was performed under the supervision of Dr. Gustavo Turecki, and the thesis contains four chapters. Chapter I is a review of the current background literature that is relevant to this thesis. Chapter II was published in *Molecular Psychiatry* and Chapter III is under review in *eLife*. Finally, Chapter IV is a summary and discussion of findings from Chapters II and III and includes concluding remarks and future directions. Tests providing links between the individual studies were written in accordance with the guidelines outlined by the McGill University Graduate and Postdoctoral Studies Office.

CONTRIBUTION OF AUTHORS

Chapter I

This chapter contains a review of the literature pertaining to the neurobiology of cocaine use disorders, from neural substrates, to neurotransmitter systems, and finally transcriptomic and epigenetic alterations. The writing and literature review were conducted by the thesis author under the supervision of Gustavo Turecki. Carl Ernst and Deborah Mash supervised the writing and review of the latter half, which reviews the evidence linking DNA methylation to cocaine-related behaviors.

Chapter II

The thesis author conceived, designed, and coordinated all aspects of the research and data analysis, as well as the interpretation of the data and the preparation of the manuscript. G.G.C was responsible for RRBS library preparation and J.Y generated the in vitro data from the 3C and dCas9-

DNMT3A experiments. The animal work was performed in collaboration with A.L., K.C.T. and B.L. under the support and guidance of E.C. and E.N. T.F, J-F.T., Z.A. and C.E provided bioinformatic and data analytical support. V.Y. provided support with the FANS experiments. Resources, including samples and support with data interpretation came from C.N, N.M, E.N, T.F. and D.C.M. G.T contributed to the design and fully supported this study, was responsible for overseeing all experiments and data, and the preparation of this manuscript. All authors discussed the results presented in this chapter.

Chapter III

The thesis author designed the study and oversaw the experiments in this chapter, as well as analysed and interpreted the data and prepared the manuscript. G.G.C. prepared the RRBS libraries and L.F. ran the qPCR experiments. G.M performed the luciferase assay and V.Y provided support with the FANS experiments. J-F.T. performed bioinformatic analyses and B.L. and E.C. performed the behavioral experiments in mice. C.N and N.M. provided support in the study design and interpretation of the data, support for the animal work was provided by E.N. and D.C.M. provided post-mortem samples. G.T. oversaw all experiments, including the interpretation of the data and the preparation of the manuscript, and all authors discussed the results presented in this chapter.

Chapter IV

This chapter contains a discussion, conclusions, and suggestions for avenues of future research. This chapter was written by the thesis author under the guidance of Gustavo Turecki.

ORIGINAL CONTRIBUTIONS TO KNOWLEDGE

In our first study, published in *Molecular Psychiatry*, we characterised the cocaine dependent methylome in the human caudate nucleus, and identified a novel cocaine-responsive regulatory element

with the Iroquois Homeobox 2 gene (*IRX2*) that regulates local gene expression and three-dimensional chromatin architecture. Not only is this the first report on the association between DNA methylation alterations and cocaine use disorders in humans, but it also identifies over 100 regions of the genome that are differentially methylated and are prime targets for further investigation. The largest cluster of CpGs that we found to be differentially methylated was within the *IRX2* gene body, and we used a combination of cell-type specific sequencing, animal models, and epigenome editing techniques to characterize its involvement in local chromatin regulation. Although *IRX2* is known to form a three-dimensional enhancer loop with its neighbor*, IRX1*, in animals, we have shown that his also occurs in human cells. Furthermore, we demonstrate that this region is sensitive to long term cocaine dependence and contains a methylation-sensitive binding domain for the CTCF protein. This work adds to the growing body of evidence that DNA methylation mediates long term neuroadaptations to cocaine and adds an important advancement to our knowledge of the relationship between DNA methylation, higher order chromatin architecture and psychiatry. The second study, in review at *eLife*, builds upon my previous work by generating the first cocaine methylome in the human nucleus accumbens and identifying a novel regulatory element within the tyrosine hydroxylase (*TH*) gene. Although TH is known to be an important component of dopaminergic functioning in the midbrain and has been associated with chronic cocaine in this region, its role in the striatum is ill-defined. Our results in FANS separated nuclei and animals, suggest that the intragenic region between exons 8 and 9 of *TH* contains a cluster of CpGs that is hypermethylated in a small population of neurons. Interestingly, TH-expressing interneurons were recently identified in the striatum and our work is the first to implicate them in cocaine neurobiology. Collectively, our findings have built upon the understanding that chronic cocaine dependence is associated with cellular and molecular adaptations by bridging the gap between human and animal literature and providing diverse avenues for future exploration.

CHAPTER I: INTRODUCTION

Substance use disorders are a major burden to public health, particularly in developed countries, and are among the list of disorders that account for 20% of years lived with disability (Bucello et al., 2010; Rehm & Shield, 2019). They are associated with increased rates of mortality though direct (i.e. overdose) and indirect means (Charlson et al., 2015). In addition, cocaine is the third most commonly used illicit drug in Canada, with 10.6% of the population having used cocaine at least once in their lifetime, and while cocaine use itself is associated with negative health outcomes, the transition from drug use, to abuse, through to disorder includes added risk of psychiatric comorbidities (Adlaf, Begin, & Sawka, 2005; Degenhardt & Hall, 2012). It is estimated that between 5% and 22% of cocaine users develop cocaine dependence within the first 12 months of their initial use, and in 2019 cocaine use disorders (CUD, formally cocaine dependence) led to more than 12 000 deaths worldwide (Global Burden of Disease Collaborative Network, 2019)

According to the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders, cocaine use disorders is characterised by the presence and severity of at least two of a number of possible symptoms, including escalating drug use, neglecting responsibilities, and repeated attempts to quit without success (American Psychiatric Association, 2013). Whereas previous editions of the diagnostic checklist distinguished between drug abuse and drug dependence, based on the presence of physiological symptoms such as withdrawal and tolerance, the most recent edition collapses physical and behavioral symptoms into single category (Hasin et al., 2013). Importantly, the presence of drug cravings is now considered to be among the possible symptoms presented for a clinical diagnosis.

1. THE NEUROBIOLOGY OF COCAINE USE DISORDER

Classically, the etiology and progression of drug use disorders has been conceptualized as progression through the three stages of the “addiction cycle”: from intoxication to withdrawal/negative affect to craving and preoccupation (Koob & Volkow, 2010). A more recent model of addiction

neurobiology includes the overarching decreases in reward circuit function and increases in stress circuits — the so called “dark side” of addiction (Koob, 2013). It is thought that the transition to compulsive drug use relieves the negative aspects of the withdrawal phase through negative reinforcement, however the “dark side” hypothesis adds that the transition includes negative urgency — the tendency to act impulsively in response to stressors (Zorrilla & Koob, 2019). Each stage of the addiction cycle, as well as the stress-related impulsivity of the dark side, are accompanied by dysregulation of an expanding list of brain nuclei.

1.1 Neuroanatomical Substrates of Addiction

Each stage of the addiction cycle is associated with specific alterations in distinct neurocircuits. The connection between the ventral tegmental area (VTA) and the striatum is instrumental to intoxication, the extended limbic system is heavily involved in negative affect and withdrawal, and the connections between the prefrontal cortex, dorsal striatum and hippocampus are highly implicated in the craving/preoccupation state (Koob & Volkow, 2010).

It is well understood that all drugs of abuse acutely activate the mesolimbic pathway, which consists of dopaminergic projections from the VTA to the nucleus accumbens (NAc), and that the increased dopamine signalling is partially responsible for the motivational effects of drug use during intoxication. For example, selective lesions in this pathway block the reinforcing effects of cocaine, and injections of psychostimulants into the NAc promotes self-administration and conditioned place preference in animals (Koob & Volkow, 2010; McGregor & Roberts, 1993). Beyond the NAc, the central nucleus of the amygdala (CeA) has also been implicated in the acute, motivational aspects of intoxication, at least partially through its connection with midbrain dopaminergic afferents, and there is also evidence that cocaine injections to the prefrontal cortex can promote self administration in a manner that parallels the NAc (Barak Caine, Heinrichs, Coffin, & Koob, 1995; McBride, Murphy, & Ikemoto, 1999; McGregor &

Roberts, 1993). Finally, the ventral pallidum, a primary downstream target of the NAc, seems to play a role in cocaine-related motivation, albeit to a lesser extent than the CeA (Robledo & Koob, 1993). Together, the increased dopaminergic signalling within the extended mesocorticolimbic structures has a strong positive motivational effect while cocaine is present at the synapse.

The second stage of the addiction cycle is negative affect and withdrawal in the absence of the drug of abuse and key components of the limbic system are heavily implicated in the maintenance of this phase. In particular, the extended amygdala (including the CeA, the bed nucleus of the stria terminalis (BNST) and a transitional zone of the medial NAc shell) are thought to combine and process inputs related to both the stress-arousal system and hedonic processing (Koob & Volkow, 2010). For example, blocking noradrenergic receptors in the BNST seems to decrease withdrawal-induced place aversion (Delfs, Zhu, Druhan, & Aston-Jones, 2000). In addition, activation of the stress response system via the hypothalamic-pituitary-adrenal axis is involved in withdrawal; the amygdala contains heightened levels of multiple stress hormones during acute drug withdrawal in animals (Koob, 2008). Finally, the activity of the infralimbic cortex seems to be involved in the extinction and inhibition of drug seeking behaviors in animals (Muller Ewald & LaLumiere, 2018). Together, the extended limbic system is involved in the second stage of the addiction cycle before prefrontal cortical regions and the striatum regain control during the preoccupation phase.

Whereas the ventral striatum (including the NAc) is instrumental in the reinforcement of drug taking during the intoxication stage, the dorsal striatum becomes important in the transition to compulsive drug use (Belin & Everitt, 2008; Everitt & Robbins, 2013). For example, severing the connection between the ventral and dorsal striatum is enough to disrupt drug seeking in animals that have established a long-lasting, compulsive phenotype, but has little impact on newly trained animals (Belin & Everitt, 2008). In human patients, stimuli that have become associated with drug intake through repeated

classically conditioning pairings can elicit dopamine spikes in the dorsal striatum that are proportionate in size to the reported intensity of the drug craving (Volkow et al., 2008).

Imaging studies have also identified addiction-related dysfunction in cognitive tasks that involve multiple subregions of the frontal cortex, including attention and cognitive flexibility (prefrontal cortex) and delayed reward/reversal learning (orbitofrontal cortex) (Aharonovich et al., 2006; Calu et al., 2007). Moreover, cocaine-dependent subjects who ingest a cocaine analog have increased activity in their orbital and medial prefrontal cortices compared to non-addicted controls, which suggests that these regions are involved in the dependent state (Volkow et al., 2005). Interestingly, the prefrontal cortex and dorsal striatum are also activated in dependent subjects upon the introduction of drug-related cues and may be involved in addiction-related changes in memory function (McClernon, Kozink, Lutz, & Rose, 2009). These studies suggest that the dorsal striatum and the extended prefrontal cortex are key components of drug craving and pre-occupation.

Maladaptive behaviors at each stage of the addiction cycle have also been associated with prefrontal cortical circuits that manage response inhibition and salience attribution tasks (Goldstein & Volkow, 2011). In general, it is thought that addiction is marked by overactivity of the reward/motivation circuity, which overcomes the control circuity of the prefrontal cortices. For example, the transition from voluntary drug use to compulsive behavior is accompanied by decreased prefrontal, in favor of increased striatal control over behavior, particularly within the dorsal striatum (caudate and putamen) (Everitt & Robbins, 2005). It has also been proposed that drug dependence creates a positive feedback loop where continued drug consumption potentiates the imbalance of striatal signalling at the expense of prefrontal pathways (Volkow, Wang, Fowler, Tomasi, & Telang, 2011). Paradoxically, fMRI studies have shown that neurocircuits that are typically hypoactive in addicted subjects compared to controls, become hyperactive in response to drug-related stimuli (Volkow, Fowler, Wang, & Swanson, 2004). This phenomenon is thought to be an important mediator of

drug cravings, as the magnitude of hyperactivity is proportionally related to self-reported desire for the drug of abuse.

Processes related to the relatively new "dark side" hypothesis have also been assigned to discrete brain areas and circuits, but studies are still needed to pinpoint their connection with other addiction-related behaviors. Cognitive psychology research has assigned many addiction-related brain areas, including the striatum, prefrontal cortex, and amygdala to aspects of negative urgency, but the insula may be an important regulatory node (Um, Whitt, Revilla, Hunton, & Cyders, 2019). In particular, activation of the insula is negatively associated with negative urgency during reward-related inhibition tasks and is particularly involved in decision making when a risky choice is made instead of a safe one (Um et al., 2019; Xue, Lu, Levin, & Bechara, 2010). In addition, a resting state MRI study identified a large-scale activation network between the orbitofrontal cortex (OFC), subgenual cingulate cortex, striatum and amygdala that was enhanced in cocaine dependent subjects and was especially active during tasks of emotional impulsivity (Contreras-Rodríguez et al., 2016). This network may be an important mediator of decision-making in patients and appears to have a nominal degree of specificity since it was identified in cocaine-dependent subjects and not those who presented with gambling addiction. Since emotional (dys)regulation is implicit in behaviors related to negative urgency, researchers have also begun to investigate brain activity in this context and have found that cocaine dependent subjects show dysfunctional activity of both cortical and limbic brain regions when executing tasks that involve appraising negative emotions (Albein-Urios et al., 2014). The overlap between brain areas implicated in negative emotional processing, and those involved in compulsivity and motivation, supports the notion that the classical addiction cycle and negative urgency likely work in tandem to alter dependence-related behaviors and brain activity.

1.2 Neurotransmitter Dysfunction

As mentioned in the previous section, all drugs of abuse increase extracellular dopamine concentrations in the NAc and these rapid changes underly the salience, motivating, and conditioning effects of drug use in multiple brain areas. As a psychostimulant, cocaine has a direct impact on dopamine signalling by binding to dopamine transporters and blocking the reuptake of extracellular dopamine (Wang, Penmatsa, & Gouaux, 2015). The resulting increased dopamine signalling in the NAc is thought to drive the motivational aspect of drug use, and extended exposure to increased dopamine may lead to compulsive drug seeking in vulnerable users, increased drug use, and physiological dependence (Volkow et al., 2008). For example, D1 dopamine receptor function is necessary to establish consistent cocaine self-administration in animals and knockout mice are resistant to developing the drug seeking behavior that is typically induced by chronic cocaine exposure (Caine et al., 2007). Similarly, in the NAc shell of rhesus macaques, signalling from both D1 and D2 dopamine receptors is needed to enhance drug self-administration (Rowlett, Platt, Yao, & Spealman, 2007).

Human brain imaging studies have corroborated the lines of evidence from animal models, and also implicate dopamine in the etiology of addicted phenotypes. In fact, the rate of change in dopamine concentrations is thought to code for distinct aspects of cognitive dysfunction, where fast, transient changes in the NAc are associated with the perception of reward (Volkow & Swanson, 2003). On the other hand, slow and stable changes in striatal dopamine signalling correlate with the perception of uncertainty and punishment, and the overall tonic level of dopamine is related to general motivational states (Schultz, 2007). Paradoxically, dopamine spikes within the NAc in response to drug related cues are higher than in response to the drug itself in addicted subjects (Volkow et al., 2011). This disparity is thought to maintain continued drug seeking behaviors in search of the anticipated reward. Moreover, increasing striatal dopamine alone is not sufficient to induce drug craving in dependent patients — dopamine spikes paired with cocaine-related cues increases craving and may act to further reinforce drug seeking (Volkow et al., 2008).

The sustained dopaminergic dysfunction that accompanies chronic drug dependence is thought to contributes to each phase of the addiction cycle. Typically, PET studies in human patients indicate two major abnormalities in dopamine signalling: decreased reward-related dopamine release in cocaine addicted subjects compared to controls, even after periods of abstinence (Volkow et al., 1997), and reduced D2 dopamine receptor-ligand binding in targets of the midbrain dopamine afferents (Volkow et al., 2006). In the striatum, decreased D2 availability may be related to increasing drug intake. In the orbitofrontal and cingulate cortices, which become hypometabolic in correlation to D2 receptor availability (Volkow et al., 1993), dopaminergic dysfunction promotes self administration, and loss of dopamine sensitivity leads to escalation of binging, decreased motivation for natural stimuli, fatigue and withdrawal (Koob & Volkow, 2010).

Other neurotransmitters have also been studied in the context of cocaine neurobiology in addition to dopamine, and many have been identified as key components of altered cellular biology and behavioral phenotypes. For example, noradrenaline signalling is involved in withdrawal related behaviors, and corticotropin releasing factor plays a role in the negative emotional states associated with abstinence (Delfs et al., 2000; X. Y. Wang, Yu, Ma, Wang, & Jiang, 2019). In non-human primates, extended cocaine self-administration also increases the density of norepinephrine transporters in the extended amygdala, and post-mortem studies have identified changes in opioid receptor expression and binding in the dorsal striatum (Beveridge, Smith, Nader, & Porrino, 2005; Hurd & Herkenham, 1993). Additionally, imaging studies have identified decreased GABAergic signalling from the ventral to the dorsal striatum, and through to the motor cortex which is thought to strengthen motor responses related to drug seeking (Volkow et al., 2011).

As a key regulator of dopamine functioning, glutamate has also been a focus of addiction neurobiology research (Wang & McGinty, 1999). In human patients, increased glutamatergic activity from the medial OFC to the NAc has been observed, and animal studies have shown that glutamatergic

synapses are critical modulators of addiction-related behaviors (Dong, 2016). Exposure to cocaine induces AMPA receptor expression at “silent” glutamatergic synapses in the NAc, which already contain NMDA receptors and rarely conduct current at resting membrane potentials. Moreover, the maturation of these synapses over time depends on the location of their afferent nuclei. Projections from the PFC recruit additional AMPA receptors while projections from the amygdala recruit calcium permeable variants of glutamate receptors that vary in their conductance (Lee et al., 2013; Ma et al., 2014). Interestingly, optogenetically reversing the maturation process in drug re-exposure experiments attenuates drug seeking behaviors regardless of the afferent location which suggests that the maturation of both types of synapses are necessary for the incubation of drug craving (Lee et al., 2013; Ma et al., 2014). Cocaine sensitization after withdrawal is also associated with increased expression and phosphorylation of distinct protein subunits of NMDA and AMPA receptors in the NAc and the PFC (Zhang et al., 2007). These mechanisms appear to be, at least partially, specific to cocaine since changes in AMPA receptor expression in the NAc are not observed when animals are sensitized after amphetamine self-administration (Nelson, Milovanovic, Wetter, Ford, & Wolf, 2009). Moreover, research has begun to explore the viability of glutamatergic modulators as treatments for cocaine use disorders (Schmidt & Pierce, 2010; Uys & LaLumiere, 2008).

1.3 Transcriptomic and Epigenomic Dysregulation

There has been a recent expansion in the number of studies seeking to identify candidate genes, networks, and molecular mediators of addiction related behavior, which have explored multiple downstream targets of the mesocorticolimbic system in animals. In the medial prefrontal cortex of cocaine-seeking rats, more than 100 genes and 70 proteins show long-term differences in expression, and numerous genes are alternatively spliced and/or differentially expressed in the rodent NAc (Feng et al., 2014; W. M. Freeman et al., 2010; Guan & Guan, 2013). Moreover, some gene expression changes

appear to be sensitive to drug dosage and are not observed after acute cocaine exposure (Feng et al., 2014). In the most extensive comparison to date, Walker et al. (2018) surveyed cocaine-induced gene expression across 6 brain regions, and throughout multiple stages of the self-administration, withdrawal, and re-exposure process. They found that there are some genes, including two protein kinases (*Lmtk3* and *Map4k2*) that are commonly dysregulated in the dorsal and ventral striatum, the VTA and the basolateral amygdala. In addition, there is a unique gene signature that correlates with a composite index of addiction-like behaviors and is activated in multiple brain regions when drug cues are present (Walker et al., 2018). Interestingly, while the list of genes that becomes differentially expressed upon re-exposure is similar between treatment groups, the magnitude of the effect seems to change depending on the animal's history with cocaine; longer periods of self-administration are associated with more profound changes.

While most studies have used brain tissue homogenates to search for dysregulated gene networks, recent work from Savell and colleagues (2020) made use of single cell transcriptomic technologies in order to disentangle the contributions of distinct cell types in the rat striatum. They identified 16 distinct cell populations in the NAc and found that subgroups of D1 and D2 dopamine receptor expressing medium spiny neurons displayed the most extensive transcriptional response to acute cocaine exposure, including increased expression of multiple transcription factor genes (i.e., *Homer1, Junb, Nr4a1 and Fosb).* This suggests that the transcriptional changes that have been observed to date are likely occurring in selected cells, perhaps those activated by distinct cocaine-associated signalling changes, rather than more broadly across all cells in the NAc.

Evidence from human tissue remains scarce, however research on targeted genes/pathways, or those using small samples of post-mortem tissue have identified diverse changes in gene expression that are evidence of altered transcriptional programming. For example, GABAergic and glutamatergic pathway genes are differentially expressed in a distinct, drug specific manner; tissue from cocaine and

alcohol abusers yield non-overlapping results (Enoch, Baghal, Yuan, & Goldman, 2013; Enoch et al., 2014; Tang, Fasulo, Mash, & Hemby, 2003). In the human VTA, dozens of genes are differentially expressed after chronic cocaine dependence, including genes related to cell death, chromatin organization and dopamine metabolism (Bannon et al., 2014). Similarly, in the hippocampus, microarray and RNAseq studies have identified differentially regulated transcription in cocaine abusers, with functions ranging from extracellular matrix organization to mitochondrial membrane functioning (Mash et al., 2007; Zhou, Yuan, Mash, & Goldman, 2011). Additionally, a microarray study in a modest sample of tissue from cocaine users showed altered transcription of genes associated with signal transduction, synaptic functioning and, myelin and oligodendrocyte functioning in the NAc (Albertson et al., 2004). Together, these results suggest that chronic cocaine dependence leads to altered transcriptional profiles, in humans and animals, that may be modulated by key upstream regulators.

The most well-studied modulators of cocaine-related gene expression changes have been a group of transcription factors that are induced by cocaine (immediate early genes, IEGs). In the NAc, gene expression that is induced by short courses of cocaine treatment is dependent on the expression of cAMP response element binding protein (CREB), whereas more chronic exposures induce ΔFosB-dependent expression changes (McClung & Nestler, 2003). CREB-related gene expression is known to be involved in drug-reward perception, potentially through altered opioid receptor expression (Carlezon et al., 1998). In addition, viral induction of CREB expression in the NAc shell of rats enhances the rewarding effects of cocaine, encourages maintenance of self-administration, and increases the probability of cocaine-primed relapse after an extended abstinence (Larson et al., 2011). In fact, CREB is one of the predicted upstream regulators of gene expression changes in the PFC, NAc and basolateral amygdala of rodents, and the VTA of non-human primates (Freeman et al., 2001; Tannu, Howell, & Hemby, 2010; Walker et al., 2018). On the other hand, since ΔFosB is a highly stable variant of FosB, it is thought to maintain lasting changes in gene expression through cocaine withdrawal through its interaction with

chromatin remodelling complexes (Nestler, 2008). ΔFosB overexpressing mice show increased sensitivity to the rewarding effects of cocaine, and reduced sensitivity to the effects of certain pro-depressive complexes (Muschamp, Nemeth, Robison, Nestler, & Carlezon, 2012).

Additionally, the early growth response 3 (Egr3) transcription factor regulates cocaine-induced transcriptional programs that include alternatively spliced transcripts and have opposing effects on behavior depending on which cell type is affected (Cates et al., 2018; Chandra et al., 2015; Penrod, Anderson, & Cowan, 2018). In each case, the transcriptional changes induced by key regulators are thought to drive the cellular and behavioral phenotypes associated with the cocaine dependence. Epigenetic mechanisms - covalent modifications to DNA and associated proteins without altering the underlying genetic sequence- may serve as the interface between chronic cocaine and transcriptional dysregulation, including priming or desensitizing genes with repeated drug exposure (Robison & Nestler, 2011).

Post translational modifications to the tails of histone proteins are a highly studied class of epigenetic modifiers that dynamically facilitate the transition between active and repressed chromatin states (Jenuwein & Allis, 2001; Nestler, 2014a). Histone acetylation, which is associated with active transcription, is one of the most studied modifications in the context of cocaine dependence and is generally increased in NAc of acute and chronically cocaine treated rodents (Kumar et al., 2005; Renthal & Nestler, 2009). Altered histone 3 (H3) and histone 4 (H4) acetylation has been detected at candidate gene promoters, (Kumar et al., 2005; Sadri-Vakili et al., 2010) and chromatin immunoprecipitation (ChIP) studies have found genome-wide H3 and H4 acetylation differences that correlate with gene expression and behavioural phenotypes (Feng et al., 2014; Renthal et al., 2009). Moreover histone acetyltransferases (HATS) and histone deacetyltransferases (HDACS) are differentially expressed in the NAc and dorsal striatum of rats that specifically self-administer cocaine (Host, Dietrich, Carouge, Aunis, & Zwiller, 2011; Malvaez et al., 2013; Romieu et al., 2008; Wang et al., 2010). Inhibition of specific

HDACs also alters the level of repressive histone methylation marks in the NAc and decreases cocaine-associated behavior (Kennedy et al., 2013)

Histone methylation is associated with transcriptional activity or repression, depending on the number of methyl groups and their associated amino acid residues (Jenuwein & Allis, 2001). H3K9me3 and H3K27me3 are negative regulators of gene transcription and display altered distribution profiles in the NAc of chronically cocaine exposed rats (Feng et al., 2014; Maze et al., 2011; Renthal et al., 2009). Furthermore, G9a, a histone methyltransferase, appears to be necessary for the establishment of cocaine-related behavioral and synaptic plasticity in mice (Maze et al., 2010). H3K4me2/3 is associated with active transcription and is altered in the NAc of chronically cocaine treated rodents, both genome-wide and in association with altered gene expression at specific loci (Caputi et al., 2014; Feng et al., 2014). A study in post mortem human hippocampal tissue revealed altered H3K4me3 levels in cocaine abusers that partially overlaps with differentially expressed genes (Zhou et al., 2011). Finally, and most recently, dopamine and serotonin have been shown to have direct effects on gene expression and drug-seeking behaviors though the histone modifications dopaminylation and serotonylation (Farrelly et al., 2019; Lepack et al., 2020). While much attention has focused on the relationship between histone modifications and cocaine-dependent phenotypes, the role of DNA methylation has remained largely underexplored (Nestler, 2014a).

2. DNA METHYLATION DYNAMICS AND COCAINE IN THE BRAIN: PROGRESS AND PROSPECTS

Kathryn Vaillancourt[1], Carl Ernst[1], Deborah Mash[2] and Gustavo Turecki[1]

[1]Department of Psychiatry, McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, Quebec, Canada

[2]Department of Neurology, University of Miami Miller School of Medicine, University of Miami, Coral Gables, Florida, USA

Published in Genes (Basel). 2017 May 12;8(5):138. Doi: 10.3390/genes8050138. Modified for this thesis.

Abstract: Cytosine modifications, including DNA methylation, are stable epigenetic marks that may translate environmental change into transcriptional regulation. Research has begun to investigate DNA methylation dynamics in relation to cocaine use disorders. Specifically, DNA methylation machinery, including methyltransferases and binding proteins, are dysregulated in brain reward pathways after chronic cocaine exposure. In addition, numerous methylome-wide and candidate promoter studies have identified differential methylation, at the nucleotide level, in rodent models of cocaine abuse and drug seeking behavior. This review highlights the current progress in the field of cocaine-related methylation and offers considerations for future research.

Keywords: DNA Methylation; Epigenetics; Cocaine; Addiction

2.1 Introduction

Substance use disorders (formerly drug dependence), including cocaine use disorder, are characterized by complex behavioural symptoms, the development of physiological tolerance, and painful withdrawal symptoms (American Psychiatric Association, 2013). Pharmacologically, cocaine is a psychostimulant that increases synaptic dopamine (Heal, Gosden, & Smith, 2014); however the behavioral complexity that accompanies the transition from casual drug use to cocaine dependence points to numerous, long lasting changes in cellular functioning. Researchers in this field have described an "addiction cycle" which consists of three behavioral/psychological states: binge or intoxication, withdrawal and negative affect, and preoccupation/drug craving (Koob & Volkow, 2010). The development of chronic drug dependence involves the progression through the addiction cycle; alongside neuroadaptive changes to important components of the mesocorticolimbic dopamine system (Figure 1).

The striatum receives direct midbrain dopaminergic input from the ventral tegmental area (VTA) and is one of the most studied brain regions in relation to cocaine neurobiology. For example, in the nucleus accumbens (NAc) which is highly implicated in the motivational aspects of drug-seeking, repeated cocaine injections leads to increases in dendritic branching and dendritic spine formation (Dumitriu et al., 2012; Robinson & Kolb, 1999). These physical changes to the synaptic machinery are associated with electrophysiological adaptation throughout the addiction cycle and diverse dysregulation of RNA transcription (Cowan, Taniguchi, & Hale, 2008; Feng et al., 2014; Otaka et al., 2013). In fact, genome-wide transcriptional changes have been found in the VTA and its projection targets, in numerous animal models of cocaine dependence and in human post-mortem tissue (Albertson et al., 2004; Bannon et al., 2015; W. M. Freeman et al., 2010; D. C. Mash et al., 2007; Zhou et al., 2011). The specificity of these changes, with distinct networks of genes being up- or down-

regulated, suggests that epigenetics, which can be defined as covalent modifications to chromatin that regulate transcription, may be involved.

Although epigenetic mechanisms are a prominent aspect of research in developmental biology, oncology, and plant biology, the intersection between epigenetics and psychiatry is relatively new. Post-translational modifications of histone tails aid in the transition between active and repressed chromatin states, and are the most well-studied epigenetic mechanisms in the context of cocaine use disorders (for review see (Nestler, 2014b)) Cytosine modifications represent another mechanism of transcriptional regulation that have begun to gain attention in the cocaine literature. The most highly studied cytosine modification in mammals is 5' methylated cytosine (5mC), which occurs primarily, although not exclusively, at cytosine-guanine dinucleotides (CpGs) (Figure 2a). At gene promoters, 5mC is linked to transcriptional repression; however increasing evidence suggests that intragenic 5mC may promote gene expression (Bird & Wolffe, 1999; Kundu & Rao, 1999; Lister et al., 2013; Maunakea et al., 2010) . In addition, 5' hydroxymethylation (5hmC), an oxidative product of active DNA demethylation, may represent a stable epigenetic mark separate from 5mC and warrants investigation in relation to cocaine use disorders (Figure 2b) (Alaghband, Bredy, & Wood, 2016; Bachman et al., 2014; J. U. Guo, Su, Zhong, Ming, & Song, 2011).

The last decade saw an increase in the number of studies investigating DNA methylation and cocaine exposure, and the purpose of this review is to summarize the findings in this field. These findings can be divided into two general categories; those related to methylation machinery, the readers, and writers of cytosine modifications, and those that identify differences in the presence or absence of the marks themselves.

2.2 Cocaine-Associated Alteration of DNA Methylation Machinery

2.2a. Pharmacological Manipulations: Impact on Drug-related Behaviors

At the broadest level, researchers have pharmacologically manipulated the DNA methylation cycle and reported diverse effects on a range of cellular and behavioral phenotypes. Overall, these studies have served to highlight the complexity of the interaction between DNA methylation and cocaine-related behaviors. For example, inhibiting DNA methyltransferase (DNMT) activity, through intracerebroventricular injections of zebularine prior to cocaine injection, delays the behavioral sensitization that is characteristic of chronic cocaine use (Anier et al., 2010). Conversely, systemic injections of the methyl supplement S-adenosylmethionine (SAM), increases sensitization to cocaine (Anier, Zharkovsky, & Kalda, 2013). It is important to note, however, that that the effects of pharmacological manipulations are region specific, as injections of DNMT inhibitor directly into the NAc, rather than brain-wide, enhances the sensitization phenotype after repeated cocaine injections (LaPlant et al., 2010).

In a similar study , Han and colleagues (2010) used 5-aza-2-deoxycytidine (5-aza), to prevent de novo DNA methylation and examine the effects on the acquisition and retrieval of cocaine conditioned place preference (CPP). In this paradigm, cocaine injections are administered in a chamber with specific contextual cues and an animal's subsequent preference for this chamber is used as a measure of the rewarding effects of cocaine. Inhibiting DNMT activity in the hippocampus of C57BL/6 mice prior to training impaired their ability to acquire CPP whereas the same manipulations in the prelimbic cortex prevented the retrieval of the conditioned memory after 24 hours delay (Han et al., 2010). Conversely, systemic injections of the methyl donor L-methionine before and throughout training, prevented the establishment of cocaine CPP altogether (Tian et al., 2012). Interestingly, L-methionine had no effect on the establishment of CPP towards morphine or food. These studies suggest that there is a complex role of de novo methylation timing during the development of drug-related behavioral change, and that this role may be distinct to the cocaine context.

Another study found that daily methionine injections during ten days of intermittent cocaine exposure reduced locomotor sensitization in rats (Wright et al., 2015). Importantly, methyl supplementation had no effect on the acute locomotor response to cocaine, which suggests that DNA methylation becomes more important as the behavioral and cellular aspects of cocaine dependence develop. To expand on this finding, the authors subjected a cohort of animals to systemic methionine injection, followed by ten days of cocaine self-administration, ten days of extinction, and a single trial of cocaine-induced reinstatement of drug seeking behavior. Although supplemental methionine had no effect on the establishment or extinction of cocaine seeking behavior, animals who had received methionine injections exhibited less drug seeking behavior given cocaine during the reinstatement trial. This blunting of the reinstatement response was not seen in methionine treated animals that underwent training for sucrose pellets self-administration which reinforces the hypothesis that DNA methylation has a more pertinent role in motivation driven behaviors to stronger, more rewarding stimuli.

Recently, Massart and colleagues (2015) used an extended withdrawal paradigm and both methylation inhibition and supplementation to determine their role in cocaine-seeking behavior, specifically in the NAc. Intra-NAc injections of RG108 in rats decreased cocaine seeking behavior up to sixty days after the last day of cocaine self-administration. Conversely, if the animals were given intra-NAc injections of SAM, they displayed significantly more drug seeking behavior, as measured by active lever presses, than control animals up to two months post-training. In both experiments, the DNA methylation cycle was modified one month after the animals underwent CPP training, in the absence of cocaine or cocaine-related cues. This is highly indicative of the continuously active role of DNA methylation in maintaining cocaine-related memories, and priming relapse-related behaviors. Together, studies that have pharmacologically manipulated the methylation cycle have shown that the role of de novo methylation in cocaine-related behaviors varies depending on the stage of behavioral acquisition, the behavior being studied, and the tissue involved.

2.2b DNA Methyltransferases: The Writers of DNA Methylation

In parallel to manipulating the DNA methylation cycle itself, researchers have studied alterations in the expression and function of DNMTs themselves, in numerous cocaine paradigms, and in multiple brain regions. These are enzymes that catalyze the conversion of cytosine to 5'-methylcytosine and, in mammals, can be divided into two major classes: de novo methylators of previously unmodified cytosine residues, and maintainers of methylation signatures through DNA replication. Traditionally, DNMT3a and 3b are classified as de novo methyltransferases while DNMT1 is a maintenance methyltransferase (Hsieh, 1999; Li, 2002; Okano, Bell, Haber, & Li, 1999; Romeo et al., 2015; Turek-Plewa J. et al, 2005), although more recent evidence has shown that all three enzymes may have de novo activity (Fatemi, Hermann, Gowher, & Jeltsch, 2002; Jeltsch & Jurkowska, 2014).

Cocaine-related DNMT expression has been primarily researched in the NAc, where the expression of DNMTs is heavily dependent on the mode of cocaine administration and the experimental time course (Table 1). For example, a single intraperitoneal injection of cocaine in mice was related to increased DNMT3a transcription after acute (1.5 hour) and extended (24 hours) withdrawal (Anier et al., 2010). DNMT3b expression was induced 24 hours after the injection, but seven consecutive days of cocaine injections had no effect on any DNMT expression in this brain area. Conversely, chronic cocaine self-administration followed by 24 hours of withdrawal results in decreased mRNA expression of DNMT1 and 3a in the NAc (LaPlant et al., 2010). If withdrawal is extended to 28 days, this pattern is reversed, with increased expression of DNMT3a. In a similar study, Wright and colleagues (2015) saw increased expression of both DNMT3 proteins, and a compensatory decrease in global DNA methylation levels, after cocaine self-administration followed by extinction and cue-reinstatement in rats. Together, these studies demonstrate that the expression of canonical de novo DNA methyltransferases is dynamically regulated in the NAc in response to cocaine-seeking behavior and withdrawal.

Expression studies outside the NAc have shown that DNMTs are sensitive to cocaine exposure in the extended mesocorticostriatal dopamine pathway. In the hippocampus, DNMT3a is induced after a short withdrawal period from chronic cocaine injections, whereas DNMT3b expression remains elevated up to 24 hours after a single dose (Anier et al., 2010). As this is the only study of methylation machinery in the hippocampus, it is difficult to speculate whether the opposing expression patterns of the two enzymes have functional significance on the development of cocaine-related pathology. In the prefrontal cortex (PFC), one study saw increased DNMT3a mRNA expression and decreased DNMT3b mRNA and protein, shortly after self-administration training in mice (Tian et al., 2012). Conversely, in rats, cocaine self-administration followed by extinction and cue reinstatement saw no change in the levels of DNMT expression, nor overall DNA methylation in this brain area in rats (Wright et al., 2015). These data point to a more transient role of methyltransferases in the PFC, where methylation changes are important during the acquisition phase of drug seeking behavior, but not during the recall of drug-related memories. Importantly, all of the aforementioned studies have examined expression in whole-tissue homogenates; however, it is increasingly important to functional changes in diverse, independent cell types given the accumulation of knowledge on cell-type specific epigenetic mechanisms (Kozlenkov et al., 2014).

To date, two groups have reported cell-type specific changes in DNMT transcription after cocaine, one in microglia and the other in striatal medium spiny neurons (MSNs). In a landmark study, Heiman and colleagues (2008) used translating ribosome affinity purification (TRAP) to isolate and transcriptionally profile dopamine receptor class 1 (D1) and class 2 (D2) expressing MSNs from the striatum of cocaine-exposed BAC transgenic mice. After chronic cocaine injections, DNMT3a transcription was induced specifically in D1-expressing MSNs and was accompanied by increased GABAergic activity of these cells in response to cocaine in culture. Notably, the authors found no changes in DNMT expression in either MSN subgroup after a single dose of cocaine. These results

emphasize the importance of cell-type specificity and suggest that the increased methyltransferase expression seen in previous studies of striatal tissue may be occurring in specific subtypes of cells.

Glial cells are an understudied population of cells to investigate cocaine epigenetics, despite their implication in human cocaine-related transcriptional dysregulation (Albertson et al., 2004; Kristiansen, Bannon, & Meador-Woodruff, 2009). In immortalized microglia from mice, even brief (3hr) exposure to cocaine in vitro is followed by a lasting increase in DNMT1 and DNMT3a protein expression (Guo et al., 2016). Similarly, moderate doses of cocaine induce DNMT1 protein expression in rat primary microglia, and levels of all three classical DNMTs are increased in the microglia of mice after chronic cocaine injections. While research continues to investigate region and cell type specificity of cocaine-related methyltransferase expression, there is also evidence to suggest that cocaine alters the expression of methylation-specific binding proteins.

2.2c Methyl-Binding Proteins: The Readers of DNA Methylation

Methyl-binding proteins (MBDs) bind to methylated segments of DNA, with varied specificity, and commonly recruit chromatin remodelling complexes to translate the signal of DNA methylation into transcriptional change (Gigek, Chen, & Smith, 2016). Although there are seven members of the MBD family, methyl-CpG binding protein 2 (MeCP2) and methyl-CpG binding protein 1 (MBD1) are the only forms that have been studied in relation to cocaine. MeCP2 is highly abundant in brain tissue and mutations within its gene play a causative role in the neurodevelopmental disorder Rett Syndrome (Skene et al., 2010; Zoghbi et al., 1999). Classically, it binds to methylated DNA and recruits histone remodelers such as deacetylases (HDACs) although recent evidence suggests that it also has affinity for hydroxymethylated cytosines (Ausió, 2016; Mellén, Ayata, Dewell, Kriaucionis, & Heintz, 2012). MBD1 is less well studied, however evidence suggests that multiple isoforms of this transcriptionally repressive protein exist, with variable affinity to densely methylated promoters (Nakao et al., 2001). Variations in

the expression and function of these proteins have begun to be examined in the context of cocaine, particularly in the dorsal striatum (DCPu), the prefrontal cortex (PFC) and the hippocampus.

The first study to report on MBDs in cocaine exposure used repeated drug injections and immunohistochemistry in rats (Cassel et al., 2006). After 10 days of cocaine injections, the number of MeCP2- as well as MBD1- positive cells in the DCPu was significantly increased. This was also true for the frontal cortex and the dentate gyrus of the hippocampus. Although the authors found no change in the number of cells expressing HDAC1 or HDAC2, two binding partners of MeCP2, they saw an overall decrease in histone H3 acetylation in all three brain areas. These results provide some of the first evidence linking methyl-binding proteins to their chromatin-remodelling effects in cocaine; however, it's important to note that these results were mirrored in a group of fluoxetine-treated animals, and therefore were not cocaine-specific.

A second study used a paradigm with differential access conditions to examine the relationship between cocaine self-administration and MeCP2 expression in rats (Im, Hollander, Bali, & Kenny, 2010). In animals given extended access (6 hours per day) to a cocaine-associated lever, the number of MeCP2 positive cells in the dorsal striatum was significantly increased. What's more, MeCP2 positive cells tended to co-localize with the neuronal marker NeuN and were not increased in animals given limited (1 hour per day) access to drug self-administration. Lentiviral knockdown of MeCP2 within the dorsal striatum decreased cocaine seeking behavior, as measured by decreased number of infusions during an extended access session, and further experiments revealed the involvement of the microRNA miR-212 in this effect. These results suggest that cocaine self-administration, particularly over an extended time period or with a higher cumulative dosage, induce the expression of MeCP2 and the activation of related neurotrophic and transcriptional pathways in the dorsal striatum.

In one of the only studies to differentiate between passive cocaine exposure and active drug-seeking, Pol Boldetto and colleagues (2014) measured MeCP2 expression in rats exposed to either passive cocaine injections or operant self-administration training. They found no changes in MeCP2 mRNA expression in either the CPu or the PFC in either condition; however, the number of MeCP-immunoreactive cells in the CPu and PFC was increased in both the passive and active cocaine conditions. In a parallel experiment, the authors tested the effects of a natural food reward and saw that MeCP induction occurred in response to active reward-seeking only. It appears that exposure to cocaine alone is sufficient to induce MeCP2 protein in the PFC and the CPu, and that active cocaine seeking does not significantly amplify this effect.

In light of increased MBD expression, Mao and colleagues (2012) sought to investigate altered protein activity in response to cocaine exposure. They measured levels of brain-specific phosphorylation of MeCP2 (pMeCP2) , which is associated with decreased repressive activity (Zhaolan Zhou et al., 2006), in rats after a single cocaine injection. When the proteins were quantified 20 minutes after the drug exposure, they saw increased MeCP2 phosphorylation in the NAc; however, it took 60 minutes to induce significant phosphorylation in the CPu. The NAc results were consistent with an earlier study that found increased immunofluorescence of pMeCP2 compared to MeCP2 two hours after cocaine injection in mice (Deng et al., 2010). Whether measured as mRNA expression, protein expression, or protein activation, the effects of cocaine exposure on MBDs mirror the effects on DNMTs, with measured increases in multiple brain regions in multiple exposure paradigms.

2.2e. Methylcytosine dioxygenases – The Modifiers of DNA Methylation

The final group of methylation-associated proteins that will be discussed here are the ten eleven translocation (TET) family of enzymes. These are a recently discovered group of enzymes involved in the oxidation of 5'-methylcytosine to 5'-hydroxymethylcytosine and its derivatives 5'-formyl and 5'-

carboxylcytosine (Ito et al., 2011; Tahiliani et al., 2009). These modified nucleotides are enriched in the brain and may represent an epigenetic regulator of transcription, distinct but related to classical DNA methylation (Gross et al., 2015; J. U. Guo et al., 2011; Lister et al., 2013).

To date, there has been one published study of TET protein expression and function in relation to cocaine. Feng and colleagues (2015) repeatedly injected mice with cocaine, and examined the expression of the TET proteins in the NAc 24 hours later. They found a down-regulation of Tet1 mRNA and protein after chronic cocaine but not Tet2 or Tet3. Using a short hairpin knockdown of Tet1 in the NAc of behaving animals, they confirmed that decreased Tet1 is related to increased cocaine CPP. Conversely, overexpressing Tet1 in the NAc decreased cocaine CPP, which suggests that this protein in the NAc is an important regulator of the behavioral response to chronic cocaine exposure. Interestingly, the downregulation of TET1 mRNA observed in their animal model was mirrored in the NAc of human cocaine abusers. These findings have important implications for the generalizability of animal work to human cocaine dependency and, together with the above changes in MBDs and DNMTs, beg the questions of whether cocaine is associated with methylation changes at the nucleotide level.

2.3. Cocaine-Associated Dysregulation of Methylation Dynamics

2.3a. Global and Methylome-wide Observations

Unlike cancer phenotypes, cocaine dependence and other psychiatric disorders typically do not show broad changes in global methylation levels (Table 2). High performance liquid chromatography (HPLC) of methylation in whole brain homogenates shows no differences between cocaine treated mice and controls (Fragou et al., 2013). A more regionally defined study found that cocaine CPP was associated with small but significant decrease in total methylated cytosines in the PFC , but not in the NAc (Tian et al., 2012). Interestingly, this effect was reversed by methionine treatment and was absent

in animals who had developed CPP towards morphine or food. More recently, Feng and colleagues (2015) used liquid chromatography followed by mass spectrometry (LC-ESI-MS/MS) to confirm the absence of overall methylation changes in the NAc of chronically cocaine-treated mice. In addition, their study showed no changes in the overall level of 5hmC in this region in response to cocaine, which suggests that if cocaine-related behaviors are associated with methylation changes, they are loci-dependent and not detectable at the global level.

To date, studies of the cocaine methylome have generally used enrichment techniques to collect methylated/hydroxymethylated DNA fragments. In one study, Massart and colleagues (2015) used an antibody raised against methylated cytosine, and a promoter array (MeDip-Array) to investigate methylation in the striatum and cocaine craving in rats. With their experimental design, the authors were able to interrogate promoter methylation patterns at multiple withdrawal time points after self-administration and after cue-induced reinstatement. Overall, the authors saw hypermethylation of promoters after extended withdrawal from cocaine, and an opposing hypomethylation of promoters when cocaine-seeking was reinstated after 30 days. Although less abundant, there appear to be distinct promoters that are hypo- or hyper-methylated one day after cocaine self-administration compared to saline yoked controls. This study also included gene-wide analyses on selected genes that had been previously implicated in cocaine use and found similar dynamic patterns of methylation change throughout the withdrawal and relapse cycle. Importantly, although there was a strong overall negative correlation between promoter methylation and transcription, very few differentially methylated loci overlapped with differentially expressed transcripts. This serves to highlight the complexity of the methylation-transcription relationship and suggests that cocaine-related methylation changes may be more stably detectable than changes in transcription.

A similar study used methyl-binding protein immunoprecipitation and high throughput sequencing (MBD-seq) to profile the methylation patterns in mouse PFC after cocaine self-

administration or injection, followed by acute or prolonged abstinence and a relapse test (Baker-Andresen et al., 2015). They found distinct methylation enrichment patterns in each condition, and differentially methylated regions (DMRs) that were associated with cocaine-seeking after prolonged abstinence. In general, DMRs were enriched in gene bodies and were more methylated in the cocaine group, although there were distinct loci that appeared less methylated. Importantly, only one of their validated DMRs, Golgb1, a Golgi-related transport protein, corresponded with changes in overall gene expression, and both measures were decreased in cocaine animals. Instead, cocaine-related DNA methylation appears to be important in regulating alternative splicing, as a number of DMRs coincided with isoform specific expression changes; an idea that fits well with findings of cocaine-related alternative splicing in other brain areas (Feng et al., 2014).

In a second PFC study, Fonteneau and colleagues (2016) trained rats to self-administer cocaine, either alone or after intra-ventricular injections of DNMT inhibitors. As was seen in mice, self-administration resulted in numerous DMRs, most of which fell within gene bodies and intergenic regions. Although roughly equal numbers of hyper- and hypo-methylated sites were seen when considering relatively small differences, the ratio was skewed towards hypermethylation when considering higher effect sizes. Using a subset of DMRs that overlapped with genes, further analysis revealed that a negative correlation between methylation and gene expression can be found only for DMRs within promoter regions and that the effect is lost when considering gene body DMRs. Interestingly, the DNMT inhibited group also showed more hyper- than hypomethylation in response to cocaine self-administration which suggests that, at least in the PFC, cocaine-related hypermethylation is more related to decreased removal of methylation rather than addition of methyl groups to new locations.

To address this issue, it is important to examine the dynamics of DNA demethylation after chronic cocaine. Although the paradigm and brain region differ from the Fonteneau (2016) study, the

Feng (2015) study described earlier (see Section 2.IV), investigated genome-wide hydroxymethylation changes in response to cocaine. Using an antibody pulldown specific to 5hmC (hMeDip-seq), this group identified over 10000 peaks of differential hydroxymethylation in the mouse NAc. As was the case with methylome studies (Baker-Andresen et al., 2015; Fonteneau et al., 2016), the majority of cocaine-induced changes in 5hmC occurred in gene bodies and intergenic regions. Combining this dataset with chromatin state information revealed that 5hmC changes were enriched at enhancer sites, and regions that frequently switched chromatin states in the cocaine or control groups (Feng et al., 2015). In addition, they found that 5hmC changes were enriched at regions flanking exon boundaries, and corresponded with isoform specific gene expression, which again suggests that cocaine-induced methylation dynamics are associated with alternative splicing events. Importantly, the authors probed the stability of these changes and found increases in 5hmC could persist up to 1 month after the last dose of cocaine.

Despite no changes in global genomic 5mC or 5hmC content, the above studies have shown that cocaine induces dynamic changes in cytosine modifications at numerous sites across the genome. Researchers typically employ gene ontology or pathway analyses in order to extract functional relevance from these genome-wide data. Although it varies between studies, cocaine-related differential methylation has been associated with pathways involved in cell morphology, neuroinflammation, kinase activity, neurotransmitter-gated ion channels and cancer (Feng et al., 2015; R. Massart et al., 2015). Unfortunately, although they are helpful in organizing data sets, these analyses are limited by the availability of known gene associations, computational predictions rather than experimental data, or unknown gene functions (Khatri & Drăghici, 2005). An alternative approach to investigating cocaine-related DNA methylation has been to explore methylation changes at specific loci.

2.3b Differentially Methylated Loci

Locus-specific studies typically overlap with genes known to have transcriptional or functional alterations during cocaine exposure (Table 2). For example, the Fos family of transcription factors have been consistently implicated in mediating the relationship between cocaine-related behaviors, epigenetic changes, and cellular adaptation (Hope, 1998). Accordingly, DNA methylation dysregulation has been studied at the promoters of two immediate early genes in the Fos family. The promoter region of FosB is depleted of methylation in the NAc shortly after cocaine exposure (Anier et al., 2010). This effect is found in mice treated acutely, with a single injection of cocaine, or repeatedly, with daily injections for seven days. Similarly, the promoter region of related gene c-Fos contains differentially methylated CpGs in the mesocorticostriatal circuitry of mice who have undergone cocaine self-administration training, extinction and drug-induced reinstatement (Wright et al., 2015). In the NAc, the average methylation of this region, and the methylation of two out of twelve CpGs in particular, is decreased in cocaine-trained animals. In the PFC, there are no overall changes between groups; however, methylation at one specific CpG is increased in the cocaine group and c-Fos expression levels are positively correlated with drug seeking behavior during reinstatement. Interestingly, most of the above methylation differences, whether hyper- or hypomethylated, were rescued by methionine injections throughout training. This echoes the effects of pharmacological manipulations on cocaine behaviors (see above section) and provides promising evidence for the utility of similar interventions to modulate the molecular and behavioral effects of cocaine dependence.

Cocaine-related methylation has also been measured at specific gene promoters of protein phosphatase 1 (PP1) subunits. Chronic cocaine has been shown to reduce the expression of PP1 catalytic subunits, and the functioning of this protein and it's repressor, dopamine- and cAMP-regulated neuronal phosphoprotein (DARPP-32), are necessary for behavioral and cellular responses to cocaine (Hiroi et al., 1999; Pol Bodetto et al., 2013; Zachariou et al., 2006). Methylation of the sequence flanking the 5' end of PP1Cβ is enriched in the CPu of rats after chronic cocaine injections, and is associated with

increased binding of MeCP2 (Pol Bodetto et al., 2013). This is paired with decreased gene transcription and which can be reversed by inhibiting de novo methyltransferases, and decreased PP1Cβ positive cells throughout the striatum and the PFC. Notably, these effects were absent after a single dose of cocaine, which is contrary the results of a similar study in the NAc (Anier et al., 2010). In this study, the authors used acute and repeated cocaine injections in mice and both an enrichment and a sequencing methodology to show increased methylation of the promoter region of the PP1c gene. Together, these studies demonstrate that cocaine related changes in PP1 expression correlate with de novo methylation of subunit promoters in multiple addiction-relevant striatal regions.

There have been three other gene-driven methylation studies in the rodent striatum; each using methylation data to infer a mechanistic relationship between chronic cocaine exposure and other molecular changes. The first study investigated CDKL5, a gene whose product is thought to interact with DNMT1 and MeCP2, after repeated cocaine injections in rats (Carouge, Host, Aunis, Zwiller, & Anglard, 2010). Chronic cocaine exposure resulted in decreased CDKL5 gene expression and CDKL5 expressing cells in the striatum. These changes coincided with increased methylation of the CDKL5 gene immediately downstream of the transcription start site (TSS). In order to relate these findings to the known increases in MeCP2 expression after chronic cocaine (Cassel et al., 2006; Im et al., 2010), the authors explored the relationship between the two proteins and found that MeCP2 binding to the CDKL5 promoter is increased after chronic cocaine, and that the expression of MeCP2 and CDKL5 is inversely related in vitro. It is unclear whether the hypermethylation at this locus is caused by de novo methylation or the absence of TET activity, but epigenetic modification of CDKL5 may be an important negative regulator of downstream methylation changes.

The second study to examine methylation changes at specific genetic loci in the rodent striatum generated their targets based on genome-wide transcriptional changes. After chronically injecting mice with cocaine, Anier and colleagues (2013) performed an expression microarray study in the NAc, and

generated lists of significantly up regulated and downregulated genes. From there, they measured DNA methylation enrichment at three selected gene promoters, which were chosen based on their potential implications for neuroplasticity. Decreased expression of the solute carrier gene *Slc17a7* and the cholecystokinin gene *Cck* coincided with increased enrichment of methylated DNA at their promoters. Similarly, increased expression of the galanin neuropeptide gene Gal corresponds with decreased methylation at the Gal promoter. Interestingly, the authors found that supplementing with repeated SAM treatment reversed the hypomethylation of Gal; however, it also counterintuitively reduced the hypermethylation of Slc17a7 as well. In addition, the authors found decreased promoter methylation of the DNMT3a and 3b genes that is concurrent with their increased expression. These data reinforce the complexity of cocaine-induced methylation changes in the striatum and suggest that methylation may act to modify behaviour on multiple regulatory levels.

Finally, the most recent study of gene-specific methylation changes in the striatum investigated changes at the brain derived neurotrophic factor (*BDNF*) locus (Tian et al., 2016). Here, the authors saw increased expression of a specific isoform of this gene, *BDNF IV*, in the NAc of mice after CPP training. Importantly, this change appears to be specific to CPP itself, as there was no difference in its expression in non-conditioned, cocaine treated animals. After bisulfite sequencing, the authors identified hypomethylation at a single CpG site within the promoter region of *BDNF IV* that overlapped with the binding site for the *C-myb* transcription factor. Despite the methylation difference occurring at a single nucleotide, the authors demonstrated increased *C-myb* binding at the BDNF IV promoter in cocaine-conditioned animals. This, along with data suggesting that BDNF overexpression increases cocaine consumption (Im et al., 2010), provides convincing evidence for a pathway between DNA methylation, neurotrophic signalling, and cocaine-related behaviors in this brain region.

Although the majority of methylation studies have been on changes occurring in bulk striatal tissue, there has been a small focus on glial cell specific alterations. In the first study, Nielsen and

colleagues (2012) trained rats to self-administer cocaine for 14 days and then examined methylation changes at promoters of white matter related genes in the corpus callosum. Although they investigated three oligodendrocyte specific genes, including the myelin-integrity related proteins *Mbp* and *Plp1*, they only saw differential methylation at the promoter of the *Sox10* gene which was significantly less methylated in the cocaine-trained animals after 1 and 30 days of forced abstinence. More recently, Guo and colleagues (2016) used microglial cell lines and FACS sorted microglia from mice to investigate methylation changes at a microRNA gene. MiR-124 may be involved in suppressing microglial activation in response to neuroinflammation (Veremeyko, Siddiqui, Sotnikov, Yung, & Ponomarev, 2013) and is upregulated in microglia in response to cocaine exposure (Guo et al., 2016). This expression change is accompanied by large increases in methylation of the pri-miR-124a-1 and pri-miR-124-2 gene promoters and supports the theory that DNA methylation plays a regulatory role over other, post-transcriptional regulators, in response to chronic cocaine.

2.3c Developmental Findings

In addition to presenting a significant health burden to adolescent and adult users, prenatal exposure to cocaine is associated with impairments in brain development and cognitive functioning that may last through the school-age years (Lambert & Bauer, 2012). At the epigenetic level, research has begun to examine the effects of prenatal cocaine on DNA methylation within the developing rodent brain. Contrary to findings in adult animals, early work in this field identified global methylation decreases in the pyramidal layer of the hippocampus of male pups exposed to prenatal cocaine, at postnatal day 3 (Novikova et al., 2008). This pattern was reversed by postnatal day 30, where cocaine-exposed animals had significantly more methylated cytosine in this cell layer. Interestingly, this reversal was accompanied by site specific changes in promoter methylation, with some loci retaining their original direction of methylation change, some reversing direction and others acquiring new methylation changes by day 30. The transition from a hypomethylated state to a hypermethylated state in this layer

of cells also coincided with an increase in de novo methyltransferase expression. A similar study showed that prenatal cocaine exposure results in increased anxiety-like behaviour and impaired spatial learning in male and female offspring into adulthood (Zhao et al., 2015). Here, prenatal cocaine exposure was also associated with hypermethylation of the paternally imprinted gene *IGF-II* and a related decrease in IGF-II protein expression in the hippocampus. Although still limited in number, these works suggest that prenatal exposure to cocaine can have long lasting repercussions on DNA methylation dynamics. It remains to be known how these changes may impact future drug seeking behavior and related neurobiology.

2.4 Future Directions

The current body of literature on DNA methylation dynamics and cocaine is relatively young, and although it has generated interesting evidence of cocaine-associated dysregulation, many questions about the mechanisms, complexity, and generalizability of these findings remain. Notably, there is a distinct lack of research using non-rodent models, let alone human samples. The only non-rodent study in this field was performed in marmoset monkeys, and found differential methylation of a single CpG within the promoter of the tachykinin receptor 3 (*TACR3*) gene, in blood (Barros et al., 2013). While valuable, it is still unclear to what extent peripheral methylation studies can be used as markers of methylation levels in the brain (Walton et al., 2016). Moreover, it is becoming increasingly clear that epigenetic modifications, including DNA methylation, vary between cell types within the same tissue (Kozlenkov et al., 2014, 2015; Mo et al., 2015). It is important that future work in this field, whether using methylome-wide or candidate loci approaches, consider the impact of cellular heterogeneity and assess cocaine-related methylation changes in distinct cell types wherever possible.

The genome-wide data thus far have indicated that the majority of cocaine-related methylation changes occur in areas outside of gene promoters (Baker-Andresen et al., 2015; Massart et al., 2015);

however, all candidate gene methylation studies have been performed in promoter regions. This discrepancy has led to gaps in our knowledge that can be addressed by methylome-wide studies using updated techniques, and extensive investigations of non-promoter methylation changes. Indeed, the traditional view of the relationship between DNA methylation and transcription is becoming more nuanced, with the inclusion of hydroxymethylation and methylated gene bodies being positively correlated with gene expression and alternative splicing (Lev Maor, Yearim, & Ast, 2015; Schübeler, 2015). In addition, cytosine methylation outside of the CpG context has recently been found to be enriched in the adult mammalian brain and is likely to regulate transcription (Guo et al., 2013; Patil, Ward, & Hesson, 2014); the roles of non-CpG methylation and cocaine action remain to be seen.

DNA methylation may also have roles in managing other epigenomic and higher-level chromatin architecture. For example, decreased DNA methylation and binding of associated methyl-binding proteins is associated with impairments in histone modifications and chromatin looping (Cedar & Bergman, 2009; Tiwari et al., 2008). Although research has developed in each domain, using the same models and paradigms of cocaine abuse, there is little in the way of relating DNA methylation profiles to other epigenetic marks. Going forward, particularly as the mechanisms underlying the relationship between these phenomena are revealed, researchers should consider the broader implications of DNA methylation on the epigenomic landscape in order to holistically understand the molecular changes associated with cocaine use and ultimately, human addiction.

2.5 Concluding Remarks

Researchers have started to investigate the relationship between DNA methylation dynamics and cocaine use disorders from multiple viewpoints. Pharmacological manipulations have shown that DNA methylation plays a dynamic role in the development of cocaine related behaviors, and, in general, DNA methyltransferases and methyl-binding proteins are induced by numerous cocaine paradigms. At

the nucleotide level, cocaine exposure is associated with diverse hyper-and hypo(hydroxy)methylation in multiple genomic contexts. The stability of methylation changes, and the growing evidence of their association with neuroadaptive changes in key addiction-relevant brain regions, makes this an exciting field of study. Future studies will undoubtedly fill gaps in our knowledge of the translatability of these findings to human patients, and the place of DNA methylation in the broader cocaine epigenome.

Acknowledgments: This work is supported by the National Institute of Drug Addiction (NIDA Grant DA033684). K.V. is supported by a Canadian Institute of Health Research Frederick Banting and Charles Best Doctoral Award.

Conflicts of Interest: The authors declare no conflict of interest.

2.6 Tables and Figures

Table 1. Cocaine-associated dysregulation of DNA methylation machinery

Ref	Species *	Tissue**	Paradigm	Withdrawal	Summary Finding
			Methyltransferases and Dioxygenases		
[37]	M	Striatum	Injection	4hr	↑ DNMT3a mRNA in D1-MSNs only after chronic injection
[22]	M	NAc; Hipp	Injection	24hr	acute injection ↑ *DNMT3a/b* mRNA in NAc and DNMT3b in Hipp repeat injection had no effect on *DNMT* expression
[24]	M	NAc	Self Admin	1.5hr	acute injection ↑ *DNMT3a* mRNA in NAc and *DNMT3b* in Hipp chronic injection ↑ *DNMT3a* mRNA in Hipp
				24hr or 28 days	Biphasic expression of *DNMT3a* mRNA (↓ 24hr but ↑ 28 days withdrawal)
			Injection	28 days	↑ DNMT3a mRNA
[26]	M	PFC	Self Admin	2hr	↑ DNMT3a mRNA and ↓ DNMT3b mRNA and protein
[55]	M	NAc	Injection	24hr	↓*TET1* mRNA and protein
	H	NAc	Post mortem	N/A	↓*TET1* mRNA
[27]	R	NAc ;PFC	Self Admin	0hr	↑ DNMT3a/b mRNA expression in NAc only
[39]	M	BV-2 cells	24hr Exposure	N/A	↑ DNMT1 and DNMT3a protein
	R	Microglia	24hr Exposure	N/A	↑ DNMT1 protein and all DNMT mRNA
	M	Microglia	Injection	1hr	↑ all DNMT mRNA
			Methyl-Binding Proteins		
[46]	R	DCPu; FC; DG	Injection	15hr	↑ MeCP2 and MBD1 positive cells in all areas
[51]	M	NAc	Injection	2hr	↑ phosphorylated MeCP2
[47]	R	DCPu; PFC; Hipp	Self Admin	24hr	↑ MeCP2 protein and positive neurons in DCPu with extended access ↓ MeCP2 protein a in PFC with extended access ↑ MeCP2 protein in Hipp with restricted or extended access
[49]	R	CPu, NAc, PFC	Injection	20 min	↑ MeCP2 phosphorylation in NAc
				60 min	↑ MeCP2 phosphorylation in CPu
[48]	R	CPu; PFC	Injection/ Self Admin	5-15hr	↑ MeCP2 expressing cells with no change in mRNA

* M=mouse; R=rat; H=human ** NAc=Nucleus Accumbens; CPu= caudate and putamen; PFC= prefrontal cortex; Hipp = hippocampus; DG= dentate gyrus

Table 2. Cocaine-associated dysregulation of methylation dynamics

Ref	Species*	Tissue**	Paradigm#	Withdrawal	Technology##	Summary Finding
			Global Methylation Level			
[26]	M	NAc; PFC	CPP	Not reported	HPLC	Slight decrease in 5mC in PFC
[56]	M	WB	Injection	1 hr	HPLC	No differences
[55]	M	NAc	Injection	24 hr	LC-ESI-MS/MS	No differences in 5mC nor 5hmC
			Cocaine Methylome			
[58]	M	PFC	Injection Self-Admin	0-2hr	MBD-seq	distinct patterns of DNA methylation after active and passive cocaine 29 persistent DMRs after self-administration (↑ 24 and ↓ 5)
[55]	M	NAc	Injection	24 hr	hMeDip-seq	5hmC alterations at enhancer regions and alternative spliced sites
[28]	R	NAc	Self Admin	1 or 30 days	MeDip-Array	Dynamic (hypo and hyper-) methylation after 1 day withdrawal Locus-specific enhancement or reversal of early methylation changes after 30 days
[59]	R	PFC	Self Admin	24 hr	MBD-seq	Cue-reinstatement reversal of many withdrawal induced changes More hyper- than hypomethyated DMRs at higher differential methylation ratios
			Differentially Methylated Loci			
[22]	M	NAc	Injection	24 hr	MeDip Bis-qPCR	↑ *PP1c* promoter methylation after acute and chronic cocaine
				1.5hr	MeDip-qPCR	↓ *fosB* promoter methylation after acute and chronic cocaine
[65]	R	Striatum	Injection	15hr	Bis-seq	↑ *Cdkl5* promoter methylation
[27]	R	NAc	Self Admin	0hr	Bis-seq	↓ *c-Fos* promoter methylation
[67]	R	CC	Self Admin	30 days	Bis-seq	↓ *Sox10* promoter methylation
[23]	M	NAc	Injection	24 hr	MeDip-qPCR	↑ *Scl17a7* and *Cck* promoter methylation and ↓ *Gal, DNMT3a and DMNT3b* promoter methylation
[48]	R	CPu	Injection	12 hr	MeDip-qPCR	↑ *PP1Cβ* promoter methylation after chronic cocaine
[39]	M	Microglia	Injection	1hr	Bis-seq	↑ *Mir124* promoter methylation
[66]	M	NAc	CPP	2hr	Bis-seq	↓ *BDNF IV* promoter at single CpG

* M=mouse; R=rat; P=Non-human primate ** NAc=Nucleus Accumbens; CPu= caudate and putamen; PFC= prefrontal cortex; WB= whole brain; CC= corpus callosum #CPP= conditioned place preference ## HPLC= high performance liquid chromatography; LC-ESI-MS/MS = liquid chromatography-electrospray ionization tandem mass spectrometry; MBD-seq = methyl-binding protein capture – sequencing; MeDIP=methylated DNA immunoprecipitation; hMeDIP=hydroxymethylated DNA immunoprecipitation; Bis-seq=sodium bisulfite conversion-sequencing; qPCR= quantitative polymerase chain reaction

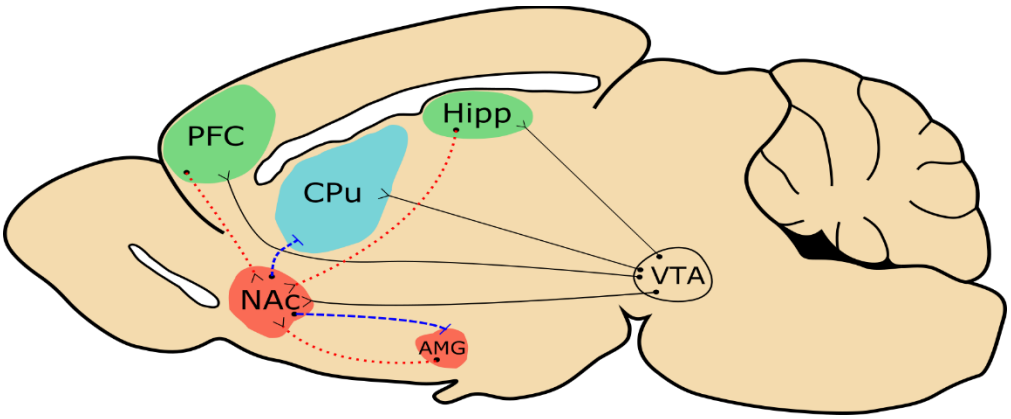


Figure 1. The simplified mesocorticolimbic pathway in the rodent brain. Solid black lines represent dopaminergic projection; dashed blue lines represent GABAergic projections and dotted red lines represent glutamatergic projections. Regions in green are implicated in the drug craving, blue in binge, and red in the withdrawal stages of the addiction cycle. PFC = prefrontal cortex; Hipp = hippocampus; CPu = caudate and putamen; NAc = nucleus accumbens; AMG = amygdala; VTA = ventral tegmental area.

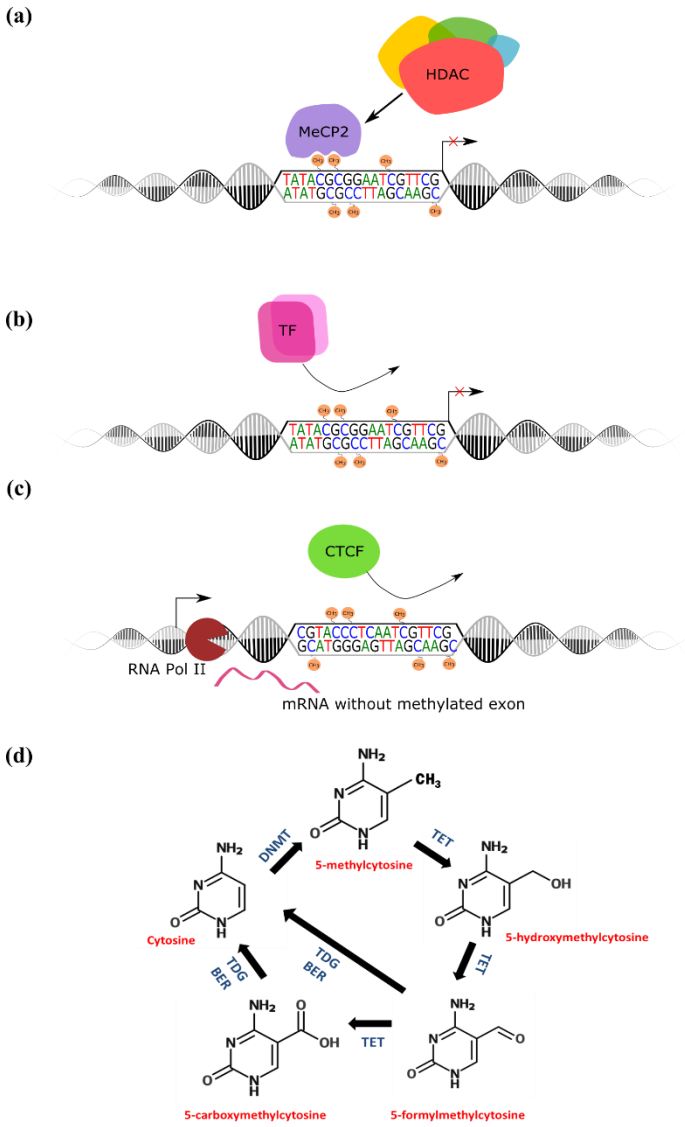


Figure 2. DNA methylation functions and dynamics. (**a**) methylated cytosines within gene promoters recruit methyl-binding proteins and chromatin remodeling complexes to prevent gene transcription; (**b**) methylated gene promoters prevent transcription factor binding; (**c**) exonic methylation regulated CTCF-mediated exon inclusion; (**d**) The cytosine modification cycle. MeCP2 = methylated-CpG binding protein 2; HDAC = histone deacetylase; DNMT = DNA methyltransferase; TET = ten-eleven translocation protein; TDG = thymine-DNA glycosylase; BER = base excision repair; TF = transcription factor; CTCF = CCCTC-binding factor; RNA Pol II = RNA polymerase II.

3. OBJECTIVES

Given the lack of information pertaining to the relationship between cocaine use disorders and DNA methylation in humans, the primary objective of this thesis project was to generate the first human cocaine methylome data using post-mortem brain samples from dependent subjects. This project sought to profile methylomic aberrations associated with disease state in two highly addiction-relevant brain areas, the nucleus accumbens and the caudate nucleus, in 25 individuals with a history of cocaine dependence and 25 drug-free and psychiatrically healthy controls, using reduced representation bisulfite sequencing (RRBS). Furthermore, we sought to identify regions of the genome that are differentially methylated (DMRs) between cases and controls in both brain areas.

The secondary objective of this work was to further investigate, and replicate interesting DMRs using targeted bisulfite amplicon sequencing, data from animal models of cocaine self-administration, and cutting-edge in vitro tools. We focused on two newly identified regulatory elements and explored their potential as genomic regulators in response to cocaine dependence.

In the following chapters, we have investigated the relationship between these novel intragenic regulatory elements, chromatin biology, and gene expression. We began by identified a group of CpGs within the third exon of the Iroquois Homeobox 2 (*IRX2*) gene that was hypomethylated in caudate neurons following cocaine dependence, and investigated its role in three-dimensional chromatin architecture, CTCF protein binding, and regulation of its own and neighboring gene expression. We conclude with the discovery of an intragenic cluster of CpGs spanning portions of the 8^{th} and 9^{th} exon of the tyrosine hydroxylase (*TH*) in the human caudate, and an exploration of it's potential as an EGR1-sensetive enhancer.

CHAPTER II: COCAINE-RELATED DNA METHYLATION IN CAUDATE NEURONS ALTERS 3D CHROMATIN STRUCTURE OF THE *IRXA* GENE CLUSTER

Cocaine-related DNA methylation in caudate neurons alters 3D chromatin structure of the *IRXA* gene cluster.

Kathryn Vaillancourt[1, 2], Jennie Yang[1], Gary G. Chen[1], Volodymyr Yerko[1], Jean-François Théroux[1], Zahia Aouabed[1], Alberto Lopez[3], Kimberly C. Thibeault[3], Erin S. Calipari[3,4], Benoit Labonté[4], Naguib Mechawar[1,2,5], Carl Ernst[5], Corina Nagy[1,2], Thierry Forne[6], Eric J. Nestler[4], Deborah C. Mash[7] and Gustavo Turecki*[1, 2, 5]

[1]McGill Group for Suicide Studies, Douglas Hospital Research Center; [2]Integrated Program in Neuroscience, McGill University; [3]Department of Pharmacology, Department of Molecular Physiology and Biophysics, Department of Psychiatry and Behavioral Sciences, Vanderbilt Center for Addiction Research; Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA; [4]Nash Family Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY USA; [5]Department of Psychiatry, McGill University, Montreal, Quebec, CA; [6]Institute de Génétique Moléculaire de Montpellier, CNRS, Université de Montpellier, FR; [7]Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, USA.

Keywords: cocaine, addiction, epigenetics, DNA methylation, 3C, 3D chromatin structure, psychiatry, post-mortem brain

Published in Molecular Psychiatry. 2020 Oct 12; Epub ahead of print. DOI: 10.1038/s41380-020-00909-x. Modified for this thesis.

Abstract: Epigenetic mechanisms, like those involving DNA methylation, are thought to mediate the relationship between chronic cocaine dependence and molecular changes in addiction-related neurocircuitry but have been understudied in human brain. We initially used reduced representation bisulfite sequencing (RRBS) to generate a methylome-wide profile of cocaine dependence in human post-mortem caudate tissue. We focused on the Iroquois Homeobox A (*IRXA)* gene cluster, where hypomethylation in exon 3 of *IRX2* in neuronal nuclei was associated with cocaine dependence. We replicated this finding in an independent cohort and found similar results in dorsal striatum from cocaine self-administering mice. Using epigenome editing and 3C assays, we demonstrated a causal relationship between methylation within the *IRX2* gene body, CTCF protein binding, 3D chromatin interaction, and gene expression. Together, these findings suggest that cocaine-related hypomethylation of *IRX2* contributes to the development and maintenance of cocaine dependence through alterations in 3D chromatin structure in the caudate nucleus.

Keywords: cocaine, addiction, epigenetics, DNA methylation, 3C, 3D chromatin structure, psychiatry, post-mortem brain

1. Introduction

Like other drug use disorders, cocaine dependence is characterized by cycles of binging, preoccupation and compulsive drug seeking behaviors despite negative outcomes (American Psychiatric Association (APA), 2000). The development and maintenance of dependence-related behaviors, in both humans and animal models, is accompanied by profound alterations in gene expression and lasting changes in cellular plasticity in the mesolimbic dopamine neurocircuitry (Albertson et al., 2004; Bannon et al., 2015; Freeman et al., 2010; Mash et al., 2007; Nestler, 2014; Zhou, Yuan, Mash, & Goldman, 2011). Accordingly, multiple targets of midbrain dopamine projections display widespread epigenetic alterations, particularly histone post translational modifications, that may mark distinct phases of dependence and withdrawal (Nestler, 2014). DNA methylation may act to stabilize dependence-related gene expression programs and it has become an active area of research in cocaine addiction neurobiology.

Cocaine-related changes in DNA methylation have primarily been measured at the level of individual gene promoters, although methylome-wide studies have begun to appear in animal models (for review, see (Vaillancourt, Ernst, Mash, & Turecki, 2017)). Furthermore, research is just beginning to investigate how changes in methylation are likely to contribute to maladaptive behavioral phenotypes (Engmann et al., 2017; LaPlant et al., 2010). However, the relationship between DNA methylation and cocaine dependence has been understudied in the human brain. Of particular interest is the caudate nucleus, as it appears to be necessary for the development of addiction related drug cravings and has been implicated in the transition from recreational drug use to dependence (Belin & Everitt, 2008; Everitt & Robbins, 2013; Garavan et al., 2000; Volkow et al., 2006). Here, we report on the findings of the first methylome-wide study of cocaine dependence in the human caudate nucleus, using post-mortem tissue samples. We also present supporting evidence for a role of cocaine-related gene body

methylation of the *IRX2*, a gene located in a region containing the largest cluster of differentially methylated CpGs from our methylome-wide analysis, in regulating local chromatin architecture and expression of two genes in the *IRXA* neurodevelopmental gene cluster.

2. Results

Chronic cocaine dependence in humans is associated with genome-wide changes in DNA methylation in the caudate nucleus

To understand how DNA methylation patterns in chronic cocaine users may differ from unaffected non-cocaine users in the caudate nucleus (Figure 1a), we performed reduced representation bisulfite sequencing (RRBS) from 25 cases who died from cocaine intoxication that had a lifetime history of cocaine dependence, but no other diagnosed psychopathology, and 25 psychiatrically healthy drug-free controls who died suddenly. RRBS allowed us to interrogate the methylation status of genome-wide loci, while enriching for CpG islands (Gu et al., 2011). The groups were matched for commonly confounding factors such as age, post-mortem interval, and tissue pH, with small effect sizes on comparison (ps > 0.1, d < 0.3; Supplemental Table 1), and sequencing statistics were not different between the groups (ps > 0.1; Supplemental Table 2). Since DNA methylation across a region of CpGs is more likely to be biologically relevant than methylation at a single nucleotide, we combined all CpGs within 50bp of another into functional regions. We detected 6712 CpG regions containing at least 2 CpGs (Supplemental Figure 1), and 173 clusters were differentially methylated between groups when correcting for ethnicity, age, smoker status and ethanol toxicology (FDR corrected $q<0.05$; Figure 1b; Supplemental Table 3).

Although we detected differentially methylated regions (DMRs) that were both hyper- and hypomethylated in the cocaine group, there were significantly more hypermethylated regions than hypomethylated (Chi Squared Goodness of Fit Tests, $\chi^2= 26.575$; $p < 0.05$, Figure 1c). Using RNA

sequencing data from the same subjects and brain nuclei, we found that transcription of the *de novo* methyltransferase DNMT3a is increased in the cocaine group (t=2.628, df=42, p=0.0120, Supplemental Figure 2a). This finding is in-line with animal studies that have shown that cocaine exposure can induce de novo methyltransferase expression in the striatum (LaPlant et al., 2010). We found no differences in expression of the two other DNA methyltransferase genes, DNMT1 and DNMT3b (ps > 0.05, Supplemental Figure 2b, c). To determine if the corrected DMRs were functionally related, we performed PANTHER gene ontology analysis of genes that either overlapped with, or were in the closest proximity to, the differential methylation signal (Thomas et al., 2003). Although we found no enrichment for genes belonging to any particular cellular component or biological process, which is likely due to unknown long-range target genes, we found a significant enrichment for genes involved in regulatory sequence-specific DNA binding and transcriptional activation (fold enrichment=4.38, FDR=0.0253).

Most of the regions, including those which were differentially methylated, mapped onto known CpG islands (Figure 1d; Supplemental Figure 3a), and although they were highly present within gene bodies (introns, exons and intron-exon boundaries), they were significantly enriched for intergenic regions (Figure 1e; Supplemental Figure 3b, Fisher's exact test $q=7.10 \times 10^{-4}$). We used data from the 15-state core model of chromatin states from the Roadmaps Epigenome Consortium (Kundaje et al., 2015), which was generated from human caudate nucleus tissues, to annotate the hypothetical chromatin status of our CpG clusters. We found that DMRs were significantly enriched for enhancers (Fisher's exact test $q=1.76 \times 10^{-10}$), regions flanking active transcription start sites (Fisher's exact test $q=2.69 \times 10^{-2}$), and weakly transcribed and quiescent regions (Fisher's exact test $q= 5.26 \times 10^{-6}$ and 1.76×10^{-7} respectively) when compared against the list of all CpG clusters (Supplemental Figure 3c, d). When we assessed the hyper- and hypomethylated DMRs separately, we found that this effect was driven by the hypermethylated loci which were enriched in the same context as the overall list (qs = $1.30\ 5 \times 10^{-4}$ - $2.46 \times 10^{-8)}$.

DNA methylation related to gene expression in cis.

To determine whether differentially methylated regions were related to transcriptomic changes in *cis*, we generated RNA-seq data from caudate tissue from the same subjects. We calculated the fold change of all genes within 5kb of a DMR and took a liberal approach to generate a list of putative DMR-gene pairs with a nominally significant expression difference (uncorrected $p < 0.1$) and statistically significant differences in methylation (q value < 0.05). This analysis identified 23 DMR-gene pairs (Table 1).

Given our analysis strategy, we rationalized that regions with more CpGs would be most likely to represent strong, biologically meaningful signals. Interestingly, the largest and third largest DMRs overlapped with members of the Iroquois Homeobox (*IRXA*) gene family that are grouped in a highly conserved cluster within vertebrate genomes and were both upregulated according to RNA-seq (Table 1). Iroquois Homeobox 2 (*IRX2*) and Iroquois Homeobox 1 (*IRX1*) are head-to-head neighbours on chromosome 5, and code for transcription factors that are involved in embryonic patterning during neural development (Matsumoto et al., 2004). Our RRBS analysis identified 21 CpGs within the third exon of *IRX2* that were 3% less methylated (Figure 2a, Table 1), and 9 CpGs within the second exon of *IRX1* that were 11% more methylated in the cocaine group (Supplemental Figure4a; Table 1).

Decreased gene body methylation of IRX2 is associated with chronic cocaine dependence

We turned to an independent cohort of dorsal caudate tissue samples from individuals with cocaine dependence (who died by causes other than cocaine overdose) and unaffected controls to replicate the methylation findings (Supplementary Table 4). Using bisulfite amplicon sequencing, we found a significant decrease in methylation of the same region within exon 3 of *IRX2* in the cocaine group compared to controls (5.5%, t=1.908, df=31.83, p=0.033, d=0.631; Figure 2b), but no differences

in methylation of the *IRX1* region (t=1.149, df=33, p=0.13, d=0.397; Supplemental Figure 4b). Based on this study, we chose to focus on *IRX2* for further analyses.

It is well known that DNA methylation patterns are cell-type specific and are particularly important in the distinct functions of neuronal and non-neuronal cell types within the central nervous system (Kozlenkov et al., 2014; Rizzardi et al., 2019). Thus, we investigated our findings in distinct populations of nuclei, separated from samples of the human caudate from our discovery cohort, using fluorescence activated nuclei sorting (FANS). We separated intact nuclei based on DRAQ5 DNA stain fluorescence, and neuronal nuclei (including D1- and D2-medium spiny neurons, as well as GABAergic and cholinergic interneurons) from non-neuronal nuclei (glial and epithelial cells) based on the nuclear marker NeuN (Supplemental Figure 5a-e). We found the cocaine-associated decrease in *IRX2* methylation to be specific to neuronal nuclei (t=1.923, df=46, p=0.03; Figure 2c). There was no group-wise difference in methylation in non-neuronal nuclei (t=0.3254, df=48, p=0.37; Supplemental Figure 6a).

Hypomethylation of a putative CTCF binding site

Expression of *IRX2*, as well as other members of the *IRXA* family, is known to be regulated by a large evolutionarily conserved group of enhancers that form a three-dimensional chromatin loop in animals (Tena et al., 2011). We hypothesized that cocaine-related methylation might interfere with this regulatory framework, perhaps by impairing the binding of CCCTC-binding factor (CTCF) —a well-studied transcriptional repressor involved in anchoring three-dimensional chromatin structures, that can be inhibited by methylation at its binding site (Hashimoto et al., 2017). We searched our target sequence for putative CTCF binding sites using its consensus sequence (Kim et al., 2007), as well as the most highly enriched motifs from chromatin immunoprecipitation data generated by the ENCODE consortium (accessed through FactorBook (Wang et al., 2012)). We found a known peak motif (3'-AGGGGGCG-5') 96

base pairs upstream of our DMR and a putative CTCF consensus sequence (3'-CCGCGGGGCGCGG-5') spanning 4 CpGs within the DMR itself. When considering the methylation state of the consensus sequence separately from the overall region, we found a main effect of cocaine status in both whole tissue homogenates from the replication cohort ($F(1, 35)=4.333$, $p=0.045$, Figure 2d)) and in neuronal nuclei ($F(1,46)=6.284$, $p=0.016$, Figure 2e)), but not in non-neuronal nuclei from the discovery cohort ($F(1,48)=0.072$, $p=0.78$, Supplemental Figure 6b)). Post hoc comparisons showed the fourth CpG to be significantly less methylated in the cocaine group in the tissue ($t=2.55$, $df=140$, $p=0.012$, Figure 2d) and in neuronal nuclei ($t=3.73$, $df=184$, $p < 0.001$, Figure 2e), but not in the non-neuronal fraction (Supplemental Figure 6b).

Next, we sought to corroborate our findings in a well-studied mouse model of self administration (Supplemental Figure 7). After removing outliers using ROUT tests (Q=2%), we found a significant effect of group ($W(2,15.7)=6.64$, $p=0.008$, Supplemental Figure 8) where the cocaine group (n=6) was significantly less methylated than the saccharin group (n=9, $p=0.029$) and nominally significantly hypomethylated compared to controls (n=14. $p=0.054$). Notably, this sequence contains the only occurrence of the canonical CTCF binding site (5'-CCGCGCCGCGCGGTGG-3') in the entire 5kB *Irx2* gene, and when we examined the methylation status of the upstream-most CpG, we again found methylation to be lower in the cocaine group compared to both control and saccharin animals ($W(2, 13.52)=10.06$, $p= 0.002$, cocaine vs. control $p=0.032$, cocaine vs. saccharin $p=0.016$, Figure 2f). Importantly, in both analyses, methylation in the control animals did not differ from those who were-trained to self-administer saccharin, which suggests that the cocaine-related hypomethylation is not generalizable to all reward-driven behaviors.

We were able to replicate our initial genome-wide significant finding of decreased methylation within *IRX2,* across sample cohorts, tissue types, and species, which is suggestive of a conserved and

functionally relevant genomic response. As such, we decided to explore the relationship between *IRX2* methylation and *IRXA* cluster gene expression in our sample set.

Intragenic IRX2 methylation is negatively associated with IRXA gene cluster gene expression

Our genome-wide analyses suggested that gene expression might be disrupted in the *IRXA* gene cluster in relation to cocaine dependence, since the expression of both *IRX1* and *IRX2* was increased according to RNA-seq (Table 1). In order to validate these findings, we used nanoString technology to count the number of *IRX1* and *IRX2* transcripts in RNA extracted from samples in our discovery cohort (n=21 cases and n=23 controls). We found significantly higher expression of *IRX2* in the cocaine group (Mann-Whitney U = 170; p= 0.019; Figure 3a), and although the increased expression of *IRX1* was not statistically significant (t=1.057, df=42; p=0.148; Figure 3b), we found the expression of the two genes to be highly positively correlated in our samples overall (r=0.622; $p < 0.0001$; Figure3c).

We next turned to *in vitro* modelling to explore the relationship between *IRX2* DNA methylation and gene expression of the *IRXA* gene cluster because homogeneous groups of cells allow better resolution than can be obtained through brain tissue homogenates. We measured endogenous methylation and expression levels in two distinct human cell lines; HEK293 kidney epithelial cells (ATCC, Virginia, US) and RENcell immortalized fetal midbrain cells (Millipore, Burlington, US). Since *IRX1* and *IRX2* are neurodevelopmental transcription factors, we hypothesized that their expression would be higher in RENcell neural progenitor cells (NPCs) compared to epithelial cells. We found this to be the case, with RENcells expressing both transcripts, whereas neither transcript was detectable in HEK293 samples (Figure 3d). There are likely multiple epigenetically relevant regulatory elements that contribute to the striking dichotomy in gene expression (alternative promoters and enhancers, for example); however, if the region within exon 3 has regulatory potential, we hypothesized that its endogenous

methylation level would differ between the cell types. Indeed, HEK293 cells were on average, 40% more methylated within this region than NPCs ($t=36.76$, $df=4$, $p < 0.0001$; Figure 3e).

To determine whether DNA methylation has any causal impact on changes in gene expression, we designed a CRISPR/Cas-9-based epigenome editing experiment in NPCs. We designed three guide RNAs (gRNAs) targeting the region within exon 3 and used a deactivated Cas9 (dCas9) enzyme fused with the active domain of a DNA methyltransferase (Vojta et al., 2016) to experimentally increase methylation and study its regulatory influence on gene expression dynamics. Cells that were transfected with the active construct were on average 4.9% more methylated than wild-type and 7.1% more methylated than the cells that were transfected with a dCas9-DNMT3a plasmid with a mutated methyltransferase domain (inactive) ($F(2,4)=16.9$, $p = 0.011$; Active vs WT $p = 0.021$, Active vs Inactive $p = 0.008$), Figure 3f). Importantly, methylation was increased across the CpGs within the CTCF binding site (6.7-8.4%, Supplemental Figure 9a) and unchanged within an amplicon in *IRX1* that was used as a control for off-target methylation ($F(2,6)=3.054$, $p=0.122$, Supplemental Figure 9b). Increased methylation of *IRX2* exon 3 resulted in a significant decrease in *IRX2* ($F(2,6)= 7.928$, $p=0.021$, Active vs. WT $p = 0.018$, Active vs. Inactive $p=0.037$)and *IRX1* ($F(2,6)= 8.417$, $p=0.018$, Active vs. WT $p = 0.020$, Active vs. Inactive $p=0.023$) gene expression (Figure 3g).

The three-dimensional chromatin structure of IRXA is associated with gene expression, and is altered by methylation of IRX2 exon 3

It has been shown in animal models that *Irx2* and *Irx1* share enhancer elements located within the intergenic region between them, and that the two genes are brought closer together to access these enhancers during transcription (Tena et al., 2011); but to date, no such regulatory loop has been identified in humans. We designed a 3C assay in human cells to detect the frequency of physical

proximity between the promoter of *IRX1* (viewpoint, Figure 4a), and the genomic region encompassing *IRX2* (test primers, Figure 4a), which are separated by over 850kB of linear genome. We found that, in both neural-progenitor and in kidney epithelial cells, the two genes are in close physical proximity more often than would be expected by chance (dotted grey line, Figure 4a). Interestingly, fragments 2 and 3, which encompass the first two-thirds of *IRX2*, are physically close to the 5' end of *IRX1* more often in NPCs, where both genes are expressed, than in cells where the genes are not expressed ($t=4.41$-6.54, $df=7$, $ps < 0.01$, Figure 4a). The cocaine-associated hypomethylation that we observed, including the putative CTCF binding site, is located within fragment 2, which is further evidence for a relationship between exon 3 methylation and 3D chromatin structure.

We next sought to understand whether methylation of the CpGs within this fragment could directly cause changes in chromatin architecture, and we again turned to epigenome-editing, this time in HEK293 cells which allowed us to transfect and grow the higher number of cells necessary for 3C (Supplemental Figure 10). The active plasmid increased the methylation of this region by 4.37% compared to the inactive plasmid and by 9.87% compared to untreated cells ($F(2,6)=15.28$, $p = 0.004$, WT vs Active $p = 0.003$, Active vs Inactive $p = 0.087$, Supplemental Figure 11a). We also found significantly more methylation in the actively transfected cells, compared to WT, when averaging across the entire CTCF binding site (9.91%, $F(2,6)=8.78$, $p=0.017$; Active vs WT $p = 0.012$, Supplemental Figure 11b). There were no significant effects of group on percent methylation of the off-target control (Supplemental Figure 11c).

In order to investigate whether these findings could be translated into alterations in long range chromatin structures, we assayed the local chromatin architecture in cells; particularly the frequency with which restriction fragments 2 and 3 interacted with the viewpoint in the *IRX1* gene. Strikingly, we found that methylation of *IRX2* exon 3 brought the interaction frequency of fragment 2, which contains

the putative CTCF binding site, down to levels near those expected by chance ($F(2,6)=71.03$, $p < 0.0001$, Active vs WT $p < 0.0001$, Active vs Inactive $p = 0.003$, Figure 4b). We found no significant effect of methylation on the interaction frequency of fragment 3, which does not contain the putative CTCF binding site, with *IRX1* (Figure 4b).

CTCF binds to IRX2 exon 3, and is disrupted by DNA methylation

Finally, to assess whether methylation of exon 3 could alter CTCF protein binding, we performed anti-CTCF ChIP-qPCR on wildtype HEK293 cells and cultures transfected with either the active or inactive dCas9-DNMT3A construct. Importantly, CTCF binding to the fragment of exon 3 containing the putative binding site was experimentally validated to be significantly higher than a non-specific IgG control ($F(1,5)=8.329$, $p = 0.034$; Supplemental Figure 11d). Furthermore, methylating the same sequence decreased CTCF binding compared to wildtype cells ($F(2,5) = 13.19$, $p=0.010$; Active vs. WT $p = 0.007$, Figure 4c). These data suggest that exon 3 may indeed contain a functional CTCF binding site that is sensitive to modest changes in cytosine methylation, such as those observed in caudate neurons after chronic cocaine dependence.

Based on the cumulation of data from human, mouse, and cell line experiments, we suggest that exon 3 of IRX2 contains a methylation sensitive CTCF binding site that is disrupted following long-term cocaine exposure and dependence (Figure 5).

3. Discussion

Our experiments show that chronic cocaine dependence in humans is associated with decreased methylation of an intragenic region of CpGs in the *IRX2* gene, which overlaps with a novel regulatory site for local gene expression and three-dimensional chromatin structure (Figure 5). This region is one of over 100 DMRs that we have identified in the human caudate nucleus. Although this is the first

methylome-wide study of cocaine use disorders using brain tissue from human patients, our work is well-aligned with two decades of studies in animals that have identified regions of both hyper- and hypomethylation in addiction relevant neural circuitry (Vaillancourt et al., 2017).

The caudate nucleus is increasingly implicated in the pathogenesis of drug use disorders as individuals transition from recreational use to compulsive drug seeking behaviors (Belin & Everitt, 2008). The neurons within the human caudate are mostly GABAergic medium spiny projection neurons (MSNs) surrounded by at least 4 distinct types of inhibitory interneurons (Bernácer, Prensa, & Giménez-Amaya, 2012; Tepper, Tecuapetla, Koós, & Ibáñez-Sandoval, 2010). Striatal MSNs can be classified into two major subgroups, D1 and D2 dopamine receptor expressing cells, which have opposing effects on drug-related behaviors, with D1-MSNs enhancing drug seeking while D2-MSNs inhibit these behaviors in animals (Lobo & Nestler, 2011). While technical limitations prevent us from discerning the contributions of individual neuronal subtypes, the separation of neuronal from non-neuronal methylation profiles presented here represents the first step towards a human cell-specific cocaine methylome. Future progress in single cell methylome technologies will undoubtedly guide deconvolution efforts on datasets such as those presented here.

Most of the work in the field, to date, has focused on methylation at specific gene promoters, but intragenic methylation and methylation at distal regulatory elements may have relevance to tissue-specific disease etiology. Indeed, the majority of DMRs identified here do not fall within annotated promoters and may disrupt other regulatory processes that contribute to addiction neurobiology. Additionally, although the DMRs in this study are not enriched for any one particular cellular component or biological process, the effects of human chronic cocaine dependence may not impact all cellular pathways equally; further research into epigenetic alterations of specific processes will be a welcome addition to this work.

Non-promoter elements are enriched for neuropsychiatric heritability factors, and levels of DNA methylation and chromatin accessibility in these regions has been shown to have brain-region specific effects on disease (Hannon, Marzi, Schalkwyk, & Mill, 2019; Rizzardi et al., 2019; Yao et al., 2015). Thus, although DNA methylation is perturbed in multiple addiction-related brain regions, the exact DMRs are likely to differ between brain nuclei. Furthermore, DNA methylation within gene bodies may directly promote gene expression, direct the use of alternative promoters, or regulate alternative splicing events (Maunakea, Chepelev, Cui, & Zhao, 2013; Maunakea et al., 2010; Yang et al., 2014). In recent years, it has become clear that these biological processes have important implications for psychiatric phenotypes overall (Gandal et al., 2018), and for cocaine dependence where it has been shown that, in the nucleus accumbens, repeated cocaine exposure can induce genome-wide alternative splicing events that are related to drug seeking behaviors in rodents (Cates et al., 2018; Feng et al., 2014).

Animal work has also identified distinct alterations in methylation and associated machinery that are related to different administration paradigms and exposure time courses. For example, although DNMT3A is initially decreased in the nucleus accumbens during cocaine withdrawal in mice, levels of the *de novo* methyltransferase becomes significantly increased after 28 days (LaPlant et al., 2010). Indeed, we found increased DNMT3A expression in the caudate nucleus of our cocaine-dependent samples. Similarly, distinct patterns of differential methylation emerge in studies using passive cocaine injection *versus* self-administration, which are related to long term behavioral changes (Baker-Andresen et al., 2015). Although we are unable to separate the effects of acute and chronic cocaine in our discovery cohort due to positive cocaine toxicology at the time of death, our replication cohort was negative for cocaine metabolites, and suggests that the findings with respect to *IRX2* are more likely linked to long-term dependence than to an acute pharmacological effect.

IRX2 is a transcriptional repressor that is highly expressed during neural development and may be related to social behavior in animals (Ahn et al., 2004; Kasper, Hebert, Aubin-Horth, & Taborsky, 2018). Its expression is known to be regulated by three-dimensional chromatin architecture, which in turn is regulated by the CTCF architectural protein (Gomez-Velazquez et al., 2017; Tena et al., 2011). Although this study is the first to report on the relationship between *IRX2* and cocaine dependence, it has already been shown that dependence-related behaviors rely on long-lasting alterations in the expression of transcription factors genes (Kelz et al., 1999; Nestler, 2005) and that genes involved in transcription and chromatin regulation are dysregulated in brain tissue from human patients (Bannon et al., 2014). Cocaine-related expression of transcription factor genes can be regulated by DNA methylation mechanics (Cannella et al., 2018), and are likely cell-type specific (Chandra et al., 2015), which is in line with what we have shown at this locus, where *IRX2* is more highly expressed in the cocaine group. Although the downstream targets of IRX2 regulation have yet to be experimentally identified, target prediction algorithms suggest that it may impact the expression of genes including *ADAM10,* a metalloprotease that has been linked to multiple psychiatric diseases and may be involved in the cognitive impairments that can accompany long term psychostimulant use (Rouillard et al., 2016; Saftig & Lichtenthaler, 2015; Shukla, Maitra, Hernandez, Govitrapong, & Vincent, 2019). In our dataset, ADAM10 expression is negatively correlated in cocaine subjects but not controls, which may be indicative of a cocaine-specific gene expression program, although more work is needed. Additionally, little is known about the dynamics of epigenetic regulation in the *IRXA* gene cluster during neuronal development — future work in animals should identify how the relationship between methylation, expression, and three-dimensional chromatin structures changes throughout development.

We have shown that cocaine-related methylation of *IRX2* exon 3 is negatively associated with gene expression through decreased frequency of three-dimensional chromatin structure. This is in line with evidence that suggests that DNA methylation can compete with CTCF binding, especially at specific

CpGs at key regulatory sites (Hashimoto et al., 2017; Maurano et al., 2015). Moreover, repeated cocaine administration has been shown to increase DNA methylation and decrease CTCF-mediated chromatin looping at the *Auts2-Caln* locus in mice (Engmann et al., 2017).

Like other work using post-mortem samples, this study presents limitations that need to be considered when interpreting the findings. First, the molecular profiles gathered from these tissues highlight the epigenetic landscape immediately prior to death, and although every effort is made to characterize the demographic information of the donors, we are unable to account for corollary factors such as lifestyle, and lifetime history of non-dependent drug exposure that could influence DNA methylation. Complementary evidence from animal models, such as has been presented here, can begin to account for the effect of extraneous factors. Similarly, although we were able to distinguish between broad categories of cell-types (neurons vs. non-neurons), the magnitude of methylation differences observed in our study suggest that the signal is coming from a relatively rare cell type and being masked by cellular diversity. Nonetheless, small changes in methylation have been shown to have physiologically relevant effects on transcription factor binding, and RNA transcription (Zhang et al., 2010, 2018), and have previously been associated with cellular and molecular alterations in post-mortem psychiatric research (Gross, Fiori, Labonté, Lopez, & Turecki, 2013; Iwata et al., 2014; Lutz et al., 2017).

Future work on the *IRX2* locus, as well as other DMRs identified in this study, should incorporate information about additional levels of epigenetic regulation, including histone modifications as well as DNA modifications outside of the canonical CpG methylation context. For example, non-CpG methylation (CpH), N6-Methyladenosine (m6A) and hydroxymethylation are epigenetic regulators that are highly abundant in the brain and are likely be important mechanisms to drug dependence (Feng et al., 2015; Guo et al., 2013; Li et al., 2019). Additionally, direct manipulation of *Irx2* in animals will allow important insight into the behavioral consequences of cocaine-related epigenetic changes. Furthermore,

although studies suggest that dependence to other psychostimulants, including amphetamine, associates with DNA methylation changes, direct comparisons between drugs of abuse, and between addiction-related brain regions, will add specificity to epigenome-wide studies. Additionally, as data from single-cell epigenomic experiments continue to become available, researchers will be able to detect differences in rare cell types that are currently masked by bulk and near-bulk tissue experiments (Luo et al., 2017).

4. Methods

Subjects

All methods used in this study were approved by the Douglas Hospital Research Ethics Board, and written informed consent was obtained from the next-of-kin for each subject. Autopsy and tissue sampling were performed in accordance with the established standards of the University of Miami Miller School of Medicine, or the Douglas-Bell Canada Brain Bank, depending on cohort source location.

Post-mortem caudate nucleus tissues from our discovery cohort were obtained from the Brain Endowment Bank at the University of Miami Miller School of Medicine (Supplemental Table 1). Samples were dissected from the dorsolateral sector of the caudate from 25 subjects who had long term histories of cocaine dependence as determined by licenced clinicians, and who died from cocaine related complications as determined by forensic pathology and brain and blood toxicology. These subjects were selected based on the absence of toxicology for illicit drugs other than cocaine and were determined to have no other psychiatric diagnoses based on medical records and the reports of next-of kin. Drug-naïve, psychiatrically healthy control subjects (n=25) were selected from accidental or natural deaths. All subjects in this cohort were male, which is reflective of the opportunistic composition of the brain samples available at autopsy. Due to the rarity of these samples, no *a priori* power analyses were performed in relation to sample sizes.

Caudate tissue from the replication cohort of 15 cases and 21 controls was obtained from the Douglas-Bell Canada Brain Bank (www.douglasbrainbank.ca). Subjects underwent a medical chart review and proxy-based interviews that were used in the characterization of substance use, which was determined through psychological autopsy by the clinical staff of the brain bank. Case status was determined based on these results, as well as toxicology at the time of death. Age, PMI, and pH did not significantly differ between groups (t=0.79-1.73, df=34, p=0.09-0.44, Supplemental Table 4). Grey matter was dissected from the left hemisphere of all samples and stored at -80°C until further processing.

Reduced representation bisulfite sequencing (RRBS)

Tissue and Library Preparation

We extracted DNA from 20mg of frozen tissue of all cases and controls from our initial cohort using Qiagen DNA MiniKits as per manufacturer's instructions. To prepare RRBS libraries, we digested 1ug of genomic DNA with *MspI* restriction enzyme, repaired the fragment ends and ligated Illumina adapters as described in elsewhere (Chen et al., 2014). Purified libraries were treated with EpiTect fast bisulfite conversion kit (QIAGEN, Cat# 59824) according to the standard protocol and indexed through PCR amplification.

Sequencing and Bioinformatic Processing

Final libraries were sequenced on the Illumina HiSeq 2000 platform at the Genome Quebec Innovation Center (Montreal, Canada) using 50bp single end sequencing, and bioinformatics processing was performed in-house, as described (Chen et al., 2014). Bisulfite conversion efficiency was determined by the ratio of T to C at the unmethylated cytosine position added during the end-repair step of library construction. Sequencing data is available under accession number GSE6364.

Differential Methylation Analysis

We defined methylation region as any CpG within 50 bp of another CpG, with no limit on the number of CpGs in a given region, but with a minimum of at least 2 CpGs using the bumphunter 3.5 package for R. For CpGs to be included in the analysis, they must have been present in at least 25 subjects from both cases and controls and have ≥5X coverage, which resulted in 270191 CpGs that went into clustering analysis. CpG regions that had a standard deviation <5% methylation across all subjects (*i.e.*, irrespective of status) were removed to avoid comparisons between stable methylation sites. For each cluster, we performed differential analysis using a general linear model with status (cocaine or control) as a fixed factor, and age, ethnicity, smoking status and ethanol toxicology as covariates. We treated CpGs independently in a given regions and used only those regions that had a Benjamini-Hochberg FDR corrected *p*-value <0.05 and which were <0.05 when calculating a single mean from all CpGs per individual.

Cluster Annotation and Enrichment Analyses

All CpG regions were annotated relative to their genomic context, their CpG island proximity, and their predicted ChromHMM chromatin state using the *annotatr 1.10.0* package in R (Cavalcante & Sartor, 2017). We calculated enrichment q values for DMRs against all CpG clusters using the *LOLA* algorithm (Sheffield & Bock, 2016).

Gene Ontological Analysis

We annotated each DMR to its nearest Refseq gene. Gene ontologies were examined with over representation tests in the gene list analysis functions of the PANTHER classification system (www.pantherdb.org). We compared the full DMR list to all human genes with respect to molecular function, biological processes and cellular components and *p*-values were calculated using Fisher's Exact tests with FDR correction.

RNA sequencing

RNA was extracted from 100mg dissections of dorsolateral caudate nucleus using RNeasy Lipid Tissue Kits (Qiagen) according to standard procedure. The RNA integrity number (RIN) for the cases was 7.9 ± 1.28 and controls was 8.4 ± 0.78 (mean ± s.d.), and aliquots of 100ng/ul of RNA were sent for library preparation and sequencing at the Broad Institute (Cambridge, MA). Libraries were prepared using a standard non-strand specific protocol (Illumina TruSeq), including poly-A selection, and multiplexed for 50bp paired end sequencing on the Illumina HiSeq 2000 platform. Sequencing data was processed as previously described (GTEx Consortium, 2015), and fold changes and *p*-values were generated by student's t-tests. Uncorrected *p*-values of <0.1 were used to create DMR-gene pairs for follow-up investigation.

nanoString gene expression validation

For count-based RNA quantifications, custom 100bp probe sequences were designed to uniquely capture the majority of transcript variants of the genes indicated in Supplemental Table 5, by nanoString Technologies (Seattle, Washington). Each probe was associated with a unique fluorescent barcode, and 20ng/ul of total RNA, from the same extraction that was used in sequencing, was run on the nCounter system under the high field of view setting at the Lady Davis Institute (Montreal, Quebec). All normalization and statistical analyses were performed with the nSolver software from nanoString technologies. Raw probe counts were normalized to 4 negative control probes and then compared between groups, using two-tailed unpaired student's *t*-tests or their non-parametric equivalent when necessary. Two samples were removed from each group due to technical failure, and statistical outliers were removed after ROUT analysis (Q=1%) resulting in 23 controls compared to 21 cases in the *IRX2* analysis and 22 cases and controls in the *IRX1* comparison.

Fluorescence Activated Nuclei Sorting

Nuclear extraction and labelling

In order to liberate intact nuclei from the caudate nucleus tissue samples, we homogenized 50mg of frozen tissue in nuclei buffer containing 10mM PIPES (pH 7.4), 10mM KCl, 2mM $MgCl_2$, 1mM DTT, 0.1% TritonX-100 and 10X Protease Inhibitor Cocktail (Sigma Aldrich, Darmstadt, Germany). Homogenates were passed through a 30% sucrose gradient in nuclei buffer in order to separate nuclei from cellular debris, then after a wash with nuclei buffer, nuclei pellets were resuspended in blocking buffer containing 0.5% bovine serum albumin in 10X normal goat serum. Each sample was co-incubated with the DNA labelling dye DRAQ5 (1:300) (ThermoFisher, Waltham, MA) and an anti-NeuN-PE antibody (1:300) (cat no. FCMAB317PE, Millipore, Darmstadt, Germany) for 60 min at room temperature, then passed through 40uM filter caps to remove any remaining cellular debris before sorting.

Nuclei Sorting

Labelled nuclear extracts were processed the BDFACSAria III platform (BD Biosciences, San Jose, CA) according to technical specifications provided by the company. We used BD FACSDIVA software (BD Biosciences, San Jose, CA) to first isolate single, intact nuclei based on DRAQ5 fluorescence at the 730/45-A filter (DRAQ5), then to sort neuronal from non-neuronal nuclei based on fluorescence detected by the 585/42 filter (PE). Sorted nuclear fractions were stored at -20°C in sheath fluid (1X PBS) until DNA extraction. On average, we isolated 180 000 NeuN+ nuclei and 492 500 NeuN- nuclei from 50mg of tissue, with roughly 37% NeuN+ in each sample. There were no differences between cases and controls in terms of total nuclei in either fraction or in the ratio of neuronal to non-neuronal nuclei captured per dissection (ps > 0.1, Supplemental Figure 3b-d).

Nuclear DNA extraction and processing

We incubated nuclear fractions with 50X protease (Qiagen, Montreal, Canada) at 56°C for at least 12 hours to ensure thorough digestion of the nuclear membranes. Liberated DNA was precipitated onto 0.2X Agencourt AMPure XP beads (Beckman Coulter, Brea, CA) after adding 20% PEG-8000 2.5M NaCl to increase the final PEG concentration to 10%. The beads were washed twice in a magnetic stand with 70% EthOH, and then DNA was eluted in 50ul MilliQ H_2O. We measured the concentration of each DNA sample using Quant-iT PicoGreen dsDNA assays (ThermoFisher, Waltham, MA) according to manufacturer specifications.

Bisulfite amplicon sequencing

DNA extraction and conversion

We obtained genomic DNA from post-mortem, homogenate tissue samples from both the discovery and replication cohorts using QIAmp DNA Mini Kits (Qiagen, Montreal, Canada) as per manufacturer specifications. Notably, DNA for the discovery cohort was extracted from the initial dissections that were used for RRBS library construction. Genomic DNA from both cohorts (2ug/sample), as well as from all sorted nuclear fractions (>100ng/sample) was converted using EpiTect Fast 96 Bisulfite Conversion Kits (Qiagen, Montreal, Canada), diluted to 150ul using MilliQ water, and stored at -20 °C.

Library Preparation

To optimize our ability to cover the desired CpGs within each amplicon, and to increase amplicon diversity for sequencing, we designed three pairs of bisulfite specific primers per DNA strand using Methyl Primer Express Software v1.0 (Applied Biosystems, CA, USA). Redundant primers were designed to be non-overlapping, to have optimal melting temperatures of 60± 2°C and to be between 18

and 24bp long in order to optimize amplification in a multiplexed reaction (primer sequences and PCR conditions in Supplementary Tables 6 and 7). We amplified each sample using 10ul reactions consisting of 5X combined primers (10uM), 3X bisulfite converted DNA and 2X KAPA HiFi HotStart Uracil+ ReadyMix (Kapa Biosystems, MA, USA). Each strand was amplified separately, and after two rounds of paramagnetic bead purification at 0.8X, amplicons from both strands were combined and amplified for 10 additional cycles to add custom primer sequences in 20 ul reactions consisting of 2.5X sample, 5X combined CS1 and CS2 primers (10uM) and 2X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, MA, USA). After an additional round of 0.8X bead purification, we indexed each sample for 10 cycles in a 20ul reaction consisting of 2.5X amplicons, 5X indexing primers (10uM) and 2X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, MA, USA). Each indexed library went through two rounds of double ended bead purification (final ratio 0.8X) to select only those fragments in the predicted range of our amplicons (400-700bp). Final library concentrations and quality control was performed on the Agilent 2200 TapeStation (Agilent Technologies, CA, USA) before samples were pooled and sequenced.

Next Generation Sequencing

We pooled libraries to a final concentration of 2nM and included a 5-10% PhiX spike-in control for each sequencing run. Final libraries were run on the Illumina MiSeq platform (Illumina, San Diego, CA) using customized 300bp paired end sequencing as described elsewhere (Chen et al., 2017). All quality control and read alignment, without removing duplicates, were performed in-house and methylation was calculated as the percent of reads containing cytosine rather than thymine at each position.

Statistical Analysis

Samples were removed for poor sequencing (less than 5X coverage of CpGs or less than 80% of CpGs covered within an amplicon) and the number of samples used in each analysis is reflected by the

respective degrees of freedom. Methylation at each position was defined as the number of reads called as cytosine, over the number total number of reads. Percent methylation was averaged across all CpGs within an amplicon and compared between groups using unpaired *t*-tests (with Welch's correction in cases of unequal variance) or Mann-Whitney U tests where groups were unlikely to be normally distributed as per Shapiro-Wilks test. Two-tailed significance tests were used except for analyses with sorted nuclei fractions and mouse samples, where a priori hypotheses allowed for one-tailed testing. Statistical outliers were removed according to ROUT analysis (Q=1%), and final sample sizes are reported in the corresponding figure legends.

For analysis of methylation at individual CpGs within CTCF binding sites, ordinary two-way ANOVAs were performed with Group and CpG as factors. Main effects of group were dissected using t tests and the Holm-Sidak method for multiple testing corrections. Data are represented as means and error bars as standard error of the means.

Mouse reward self-administration

Jugular catheter implantation:

Animal experiments were approved by the Animal Care and Use Committee at Icahn School of Medicine at Mount Sinai. Male c57BL/6J mice, aged 6-8 weeks at the beginning of the experiment, were randomly assigned to groups, anesthetized with ketamine (100mg/kg) and xylazine (10mg/kg), implanted with chronic indwelling jugular catheters, and trained for i.v. self-administration as previously described (Johnson et al., 2019). The catheter tubing was passed subcutaneously from the back to the jugular vein and 1.0 cm of tubing was inserted into the vein and secured with silk suture. Catheters were flushed daily with ampicillin (0.5mg/kg) and heparin (10U/mL) solution in sterile saline (0.9% NaCl). Mice recovered >3d before commencing behavioral training. All animals were maintained on a reverse light

cycle (7:00am lights off; 7:00 pm lights on) and behavioral training was conducted during the animal's dark cycle.

Cocaine self-administration

Mice (n=6) were trained to self-administer cocaine as previously described. Briefly, mice were maintained at ~90% of their free-feeding weight and trained in standard mouse operant chambers (Med Associates, St Albans, USA) equipped with a white noise generator and two illuminated nose-pokes. Start of daily sessions (2h) were signaled by white noise. For each task, one nose-poke was designated as the "active poke" and the other designated as the "inactive poke". Active nose pokes resulted in a cocaine infusion (0.5mg/kg/inj, 3 sec) with a concurrent presentation of the Active nose poke light for 5 seconds. Inactive pokes resulted in no programmed consequences. Mice were trained to self-administer cocaine for 10 consecutive days under a fixed-ratio (FR1) schedule of reinforcement. For cocaine self-administering animals, acquisition (Day 1) was counted when the animal reached 70% responding on the active lever and 10 or more responses. Control animals underwent the same experimental procedures but had access to a saline-paired nose poke (n=7). For both the cocaine and saccharin self-administration groups, most mice reached acquisition criteria on the first day, and since the experimental animals were trained for 10 days, all saline animals underwent 10 days of saline self-administration as controls.

Saccharin Self-Administration

8-week-old c57BL/6J mice were ordered from the Jackson Laboratory and were housed in a 12-hour 6:00/6:00 reverse dark/light cycle. Saccharin self-administration was run during the animal's dark cycle and mice were food restricted to ~90% of free-feeding weight with water provided ad libitum. Mice (n=9) were trained to self-administer saccharin (0.1% solution in water, ~80uL/infusion) or water (n=7) for 1 h over 10 consecutive days. Briefly, activation of white noise signaled the initiation of the (1h) daily

session. In each task, one nose-poke was designated the "active" while the other was designated the "inactive". Active responses initiated a 1 sec saccharin (or water) delivery into an accessible dipper with a concurrent 5 sec presentation of both the nose poke light and dipper light under a Fixed-Ratio 1 (FR1) schedule of reinforcement. Inactive nose pokes had no programmed consequences but were recorded throughout all behavioral sessions.

Tissue preparation

After the 10th test session, animals were euthanized, and brain tissue was removed, and flash frozen at -80°C before the caudate-putamen was dissected.

Library Preparation

After standard extraction using, 500ng of genomic DNA from each sample was bisulfite converted using the EZ DNA Methylation Gold kit (Zymo Research, Irvine, California) by a blinded experimenter. Bisulfite specific primers were designed as above, using genomic regions homologous to the hg19 coordinates in the mouse genome (mm10) (Supplemental Table 6). Bisulfite DNA was subjected to the same amplification (Supplemental Table 7) and purification methods as described above, and the final purified libraries were spiked into a 300bp paired end sequencing run containing customized sequencing primers as above. After de-multiplexing and adapter trimming, sequencing reads were aligned to the mm10 mouse genome, and percent methylation at each position was determined as above.

Cell Culture

Human embryonic kidney cells (HEK293) were maintained in eagle's minimum essential medium (EMEM; ATCC, Virginia, US) supplemented with 10% fetal bovine serum (FBS; Gibco Laboratories,

Gaithersburg, MD), 100 I.U./mL penicillin and 100ug/mL streptomycin. Cells were cultured at 37°C in a humidified incubator with 5% CO_2.

Human neural progenitor cells (ReNcell; Millipore, Burlington, US) were maintained in STEMdiff neural progenitor medium (STEMCELL Technologies, Vancouver, Canada) supplemented with 100 I.U./mL penicillin and 100ug/mL streptomycin. All cell lines were tested for mycoplasma contamination and authenticated by their respective manufacturers.

Quantitative PCR

The expression of *IRX1* and *IRX2* in all *in vitro* experiments was determined using quantitative reverse transcription PCR (RT-qPCR). We used pre-designed probe-based assays for *IRX1* (Hs.PT.58.2400) and *IRX2* (Hs.PT.58.24473971) with FAM-TAMRA dyes (PrimeTime, IDT, Hampton, New Hampshire, Supplemental Table 8), and ran 10ul assays on an Applied Biosystems QuantStudio 6 instrument under default cycling conditions (ThermoFisher, Hampton, New Hampshire). Amplification curves were normalized using QuantStudio Real-Time PCR Software, and the average expression between treatment conditions was compared with two-tailed t tests in GraphPad Prism 6 (www.graphpad.com).

dCas9 epigenome editing

Plasmid preparation

The pdCas9-DNMT3A-PuroR and pdCas9-DNMT3A-Puro (ANV) plasmids were a gift from Vlatka Zoldoš (Addgene plasmids #71667 and #716840). Guide RNA (sgRNA) sequences targeting human *IRX2* exon 3 were designed using an online tool (CRISPR gRNA Design Tool, www.ATUM.bio) and purchased from Integrated DNA Technologies (IDT, Iowa, US). All 3 sgRNAs were annealed and cloned into the

expression plasmids through the *BbsI* restriction site. All cloned constructs were confirmed by Sanger sequencing using the U6 sequencing primer (Supplemental Table 9).

Transfection

HEK293 cells were seeded in 10cm culture dishes and transfected the next day, at over 80% confluence, using TransIT-293 transfection reagent (Mirus, Brampton, ON). Transfections were done with a pool of 15ug dCas9-DNMT3A plasmids carrying all three sgRNA targeting *IRX2*. The experiment was performed in triplicate and constructs carrying inactive DNMT3A were used as negative controls. 48h after transfection, cells were selected with 1.6ug/mL puromycin (Gibco Laboratories, Gaithersburg, MD) for another 48h. Cells were harvested after seven days transfection for DNA and RNA extraction (see above) as well as the nuclei preparation (see below).

RENcells were transfected by electroporation using the Neon transfection system (Invitrogen, Carlsbad, CA), according to the manufacturer's standard procedure. Briefly, cells were washed in PBS, detached from the culture vessel using accutase, pelleted by centrifugation and resuspended in Resuspension Buffer R at a final density of 13 x 10^6 cells/mL. Cells were immediately electroporated three times using 100uL neon tips at voltage 1300, width 20, and pulse 3, giving final 15ug DNA and 4 x 10^6 cells per 10cm culture dish. After electroporation, cells were immediately transferred into the prepared 10cm culture dish containing prewarmed medium, but without antibiotics. Cells were selected with 0.3ug/mL puromycin with the same time period as in HEK293 and harvested at Day 7. Percent methylation was assessed with bisulfite amplicon sequencing, and gene expression was assessed by qPCR as described above.

Statistical Analyses

Percent methylation, gene expression and interaction frequency were compared between conditions using one-way ANOVAs with Tukey's post-hoc comparisons.

Chromatin Immunoprecipitation (ChIP-qPCR)

Chromatin Preparation

HEK293 cells were grown, transfected, and nuclei were harvested as above. Aliquots of 5 million nuclei each were prepared from wildtype cells, cells transfected with the active dCas9-DNMT3A + gRNA construct (active group), and cells that were transfected with the mutant dCas9-DNMT3A construct (inactive group). Pelleted nuclei were resuspended with 1mL filtered 1X PBS and 200X Protease Inhibitor Cocktail, and cross-linked with 1% formaldehyde for 7 minutes, at room temperature, with gentle rotation. After the fixation reaction was stopped with 10X glycine, nuclei were pelleted and washed twice with 1XPBS+PIC and resuspended in Sonication Nuclear Lysis Buffer, and chromatin was prepared as per the manufacturer's protocol (SimpleChIP Plus Sonication Chromatin IP Kit, Cell Signalling Technologies, Massachusetts). Chromatin was sheared on a S220 focused ultrasonicator (Covaris, Massachusetts) under the following conditions: 150 peak power, 200 cycles/burst, duty factor 10, for 15 minutes.

Immunoprecipitation

Diluted "input" chromatin was removed before immunoprecipitation, and anti-IgG was added to chromatin from 1 x106 nuclei as per manufacturer protocol (SimpleChIP Plus Sonication Chromatin IP Kit, Cell Signalling Technologies, Massachusetts). Anti-CTCF antibody (1:25 dilution, cat no. 2899, Cell Signalling Technologies, Massachusetts) was added to the sonicated chromatin of 4 x 10^6 nuclei and both anti-IgG and anti-CTCF preparations were precipitated overnight at 4°C with rotation. Chromatin-

antibody complexes were separated with magnetic protein G beads, eluted, and de-crosslinked by proteinase K digestion at 65°C overnight (16 hours).

qPCR

DNA from anti-CTCF, anti-IgG and input fractions was purified per manufacturer's instructions (SimpleChIP® DNA Purification Buffers and Spin Columns Kit, Cell Signalling Technologies, Massachusetts), 10ul SYBRGreen qPCR reactions were prepared using the primers provided in Supplementary Table 10, and 3 technical replicates were amplified using an Applied Biosystems QuantStudio 6 instrument under default cycling conditions (ThermoFisher, Hampton, New Hampshire).

Data normalization and analysis

Ct values were averaged across technical replicates and the amplicon enrichment was calculated with the following formula: $E^{\wedge(Ct_{input}-Ct_{IP})}$ where E= log(efficiency) of each primer pair. These data were then multiplied by the input dilution factor to obtain "percent of input". For each sample, the percent input for the test amplicon was normalized to the percent input generated from the amplification of a control amplicon, located upstream of the exon 3 target, by making a ratio. These values were then analysed using Kruskal-Walis test (one outlier was removed from the wildtype group for having a 10-fold difference in normalized percent input) followed by Dunn's multiple comparisons test, comparing the wildtype and inactive groups to the actively methylated group.

Chromatin Conformation Capture (3C-qPCR)

3C library preparation

In order to assess three-dimensional chromatin structure, we used chromatin conformation capture, followed by quantitative PCR (3C-qPCR), using the protocol from Ea and colleagues (Ea, Court, & Forne, 2017). Nuclei were extracted from 5-10 million cells using a sucrose gradient, crosslinked in 1%

formaldehyde for 10 minutes, and quenched in 125mM glycine before centrifugation and resuspension in a restriction enzyme compatible 3C buffer. Samples were digested overnight at 37°C with 450U of high concentration EcoRI (Promega, Wisconsin) shaking at 200rpm, and then diluted in 4ml of ligation buffer to promote intramolecular ligation and prevent chromatin tangles. We ligated the samples with 195U high concentration T4 ligase (Promega, Wisconsin) overnight at 16°C, and then extracted the ligation products using standard proteinase K digestion, phenol chloroform extraction, and ethanol precipitation with the addition of 1ul glycogen. In order to prevent circularization or coiling of ligation products, we performed a complementary digestion using 100U of BglI (ThermoFisher, Waltham, MA), followed by phenol-chloroform extraction and ethanol precipitation of the final 3C libraries.

3C library quality control

Digestion efficiency can have a crucial impact on the outcome of 3C-based assays and as such, we assessed the digestion efficiency of each cut site within our experiment by comparison with an undigested control (UND) taken from cross linked chromatin, and a digested control (DIG) taken after EcoRI digestion but before ligation. We purified the UND and DIG control samples using proteinase K digestion, phenol chloroform extraction and ethanol precipitation, and then performed a 2-hour BglI digestion at 37°C. Using primers designed to span across each restriction site (R; Supplemental Table 11), we performed SybrGreen qPCR on the UND and DIG fractions for each sample (PowerUp SYBR Green Master Mix, ThermoFisher, Waltham, MA). To control for differences in the amount of starting material, we also amplified each fraction using primers designed for region within *GAPDH* that does not contain an EcoRI, nor a BglI cut site (C, Supplemental Table 11). We calculated the restriction digestion efficiency for each restriction site using the following formula: % Efficiency = 100 – 100/2^((Ct_R-Ct_C)$_{DIG}$-(Ct_R-Ct_C)$_{UND}$); we excluded all samples whose efficiency, averaged across all restriction sites, was less than 70%.

PCR Control Template Library

In order to determine the minimum concentration of 3C library needed for each qPCR reaction, and to compare relative interaction frequencies between primer pairs, we generated a control template library containing all possible ligation fragments, across our region of interest, in equimolar concentrations. To do this, we obtained two human BACS (RP1182M24 and RP11596I24; ThermoFisher, Waltham, MA) and combined them in equimolar concentration. We digested the BAC pool with EcoRI for two and a half hours at 37°C, and then ligated with T4 ligase overnight at 16°C. After phenol chloroform extraction, we performed BglI digestion at 37°C for two hours and then purified with phenol chloroform and ethanol precipitation for a final time. We made five-fold serial dilutions of this template library in 25ng/ul BglI digested gDNA, starting at a concentration of 25ng/ul to mimic the behavior of the 3C libraries. For each primer pair in our assay, we performed qPCR using these serial dilutions (see below) and obtained a standard curve with a slope (b) and intercept (a) that were used to normalize the Ct values of our samples.

3C-qPCR

Prior to measuring the amount of each ligation product in our 3C libraries, we measured the concentration of each library by SybrGreen PCR using primers for the non-digested site within *GAPDH*. We then adjusted the concentration of our libraries to 25ng/ul and re-measured the concentration to ensure accuracy. The final concentration values were used as loading control values during normalization.

We designed our qPCR assay to cover the ~1Mb region encompassing both *IRX2* and *IRX1* (chr5: 2744845-2752662, primers in Supplemental Table 10). We designed a constant reverse primer on the 220th EcoRI digestion fragment, which overlaps the promoter region and TSS of *IRX1*, as determined by in silico digestion. This fragment also bound to a custom PrimeTime Probe (Integrated DNA

Technologies, Iowa) that was positioned between the constant primer and the EcoRI cut site and had a 5' FAM fluorescent dye and a 3' TAMRA quencher. We designed 12 test primers, adjacent to the cut sites of fragments concentrated around both *IRX* genes (Supplemental Table 11), such that the resulting amplicons would be between 100bp and 150bp when paired with the constant reverse primer. We determined the Ct value for each ligation product in 10ul reactions run in triplicate, using 2X TaqMan Master Mix, on a QuantStudio 6 instrument using QuantStudio Real-Time PCR Software (ThermoFisher, Hampton, New Hampshire).

Data Normalization and Analysis

To account for differences in efficiency between primer pairs, we first normalized our experimental Ct values to the standard curve obtained for each reaction using our control template library as follows: Normalized Ct_1 = 10^((Ct-b)/a). In order to account for variation in the amount of template, we normalized each Ct_1 to the concentration of the input library (Ct_2 = Ct_1/loading control), and these values were used as the relative interaction frequency between each fragment and the constant. Finally, given that the ends of each hybrid ligation fragment originate from the same DNA molecule, we calculated the basal interaction level (BIL) using the procedure as previously described (Braem et al., 2008) and defined it as the relative frequency of interaction that would be expected by chance ("random collisions"). Using the BIL for each library, and the standard error of the mean of these values, we determined the "noise band" where any observed interaction within this range would be attributed to chance and not a biologically meaningful effect.

We used student's t tests to compare the average relative interaction frequencies between groups for each fragment and used the Holm-Sidak method to correct for multiple comparisons across each 3C experiment.

5. Acknowledgments

We are deeply grateful to the families of the subjects used in this study. We would also like to acknowledge the teams of technicians at the Miami Brain Endowment Bank™ and the Douglas Bell Canada Brain Bank, and the bioinformaticians who have worked on these data (A.D., R.P., and A.B.).

This work was supported by a Canadian Institute of Health Research Doctoral Fellowship awarded to K.V., National Institute of Drug Abuse Grants DA033684 awarded to G.T and D.M., R00 DA04211 to ESC, and P01 DA047233 to E.N., and grants from the Whitehall Foundation, the Edward Mallinckrodt Jr, Foundation, and the Brain and Behavior Research Foundation to ESC.

Author Contributions

Manuscript preparation: K.V., Experimental design and data collection: K.V., J.Y., G.G.C., Data Analysis: K.V., C.E., T.F., J-F.T., Z.A., Animal experiments: A.L, K.C.T., B.L, Resources and support: E.N., E.C., C.N., D.C.M., G.T.

Competing Interests

The authors declare no competing interests.

6. References

Ahn, J.-I., Lee, K.-H., Shin, D.-M., Shim, J.-W., Lee, J.-S., Chang, S. Y., … Lee, Y.-S. (2004). Comprehensive transcriptome analysis of differentiation of embryonic stem cells into midbrain and hindbrain neurons. *Developmental Biology*, *265*(2), 491–501. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14732407

Albertson, D., Pruetz, B., Schmidt, C., Kuhn, D., Kapatos, G., & Bannon, M. (2004). Gene expression profile of the nucleus accumbens of human cocaine abusers: evidence for dysregulation of myelin. *Journal of Neurochemistry*, *88*(5), 1211–1219. https://doi.org/10.1046/j.1471-4159.2003.02247.x

American Psychiatric Association (APA). (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., T). Washington:DC: American Psychiatric Association (APA).

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

Bannon, M. J., Johnson, M. M., Michelhaugh, S. K., Hartley, Z. J., Halter, S. D., David, J. a, … 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 : Official Publication of the American College of Neuropsychopharmacology*, *39*(9), 1–9. https://doi.org/10.1038/npp.2014.70

Bannon, M. J., Savonen, C. L., Jia, H., Dachet, F., Halter, S. D., Schmidt, C. J., … Kapatos, G. (2015). Identification of long noncoding RNAs dysregulated in the midbrain of human cocaine abusers. *Journal of Neurochemistry*, *135*, 50–59. https://doi.org/10.1111/jnc.13255

Belin, D., & Everitt, B. (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/10.1016/j.neuron.2007.12.019

Bernácer, J., Prensa, L., & Giménez-Amaya, J. M. (2012). Distribution of GABAergic interneurons and dopaminergic cells in the functional territories of the human striatum. *PLoS ONE*, *7*(1). https://doi.org/10.1371/journal.pone.0030504

Braem, C., Recolin, B., Rancourt, R. C., Angiolini, C., Barthès, P., Branchu, P., … Forné, T. (2008). Genomic matrix attachment region and chromosome conformation capture quantitative real time PCR assays identify novel putative regulatory elements at the imprinted Dlk1/Gtl2 locus. *The Journal of Biological Chemistry*, *283*(27), 18612–18620. https://doi.org/10.1074/jbc.M801883200

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. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *38*(34), 7516–7528. https://doi.org/10.1523/JNEUROSCI.0600-18.2018

Cates, H. M., Heller, E. A., Lardner, C. K., Purushothaman, I., Peña, C. J., Walker, D. M., … Nestler, E. J. (2018). Transcription Factor E2F3a in Nucleus Accumbens Affects Cocaine Action via Transcription and Alternative Splicing. *Biological Psychiatry*, *84*(3), 167–179.

https://doi.org/10.1016/j.biopsych.2017.11.027

Cavalcante, R. G., & Sartor, M. A. (2017). Annotatr: Genomic regions in context. *Bioinformatics*, *33*(15), 2381–2383. https://doi.org/10.1093/bioinformatics/btx183

Chandra, R., Francis, T. C., Konkalmatt, P., Amgalan, A., Gancarz, A. M., Dietz, D. M., & Lobo, M. K. (2015). Opposing Role for Egr3 in Nucleus Accumbens Cell Subtypes in Cocaine Action. *Journal of Neuroscience*, *35*(20), 7927–7937. https://doi.org/10.1523/JNEUROSCI.0548-15.2015

Chen, G. G., Diallo, A. B., Poujol, R., Nagy, C., Staffa, A., Vaillancourt, K., … Ernst, C. (2014). BisQC: an operational pipeline for multiplexed bisulfite sequencing. *BMC Genomics*, *15*, 290. https://doi.org/10.1186/1471-2164-15-290

Chen, G. G., Gross, J. A., Lutz, P.-E., Vaillancourt, K., Maussion, G., Bramoulle, A., … Ernst, C. (2017). Medium throughput bisulfite sequencing for accurate detection of 5-methylcytosine and 5-hydroxymethylcytosine. *BMC Genomics*, *18*(1), 96. https://doi.org/10.1186/s12864-017-3489-9

Ea, V., Court, F., & Forne, T. (2017). Quantitative Analysis of Intra-chromosomal Contacts: The 3C-qPCR Method. *Methods in Molecular Biology*, *1589*, 75–88.

Engmann, O., Labonté, B., Mitchell, A., Bashtrykov, P., Calipari, E. S., Rosenbluh, C., … Nestler, E. J. (2017). Cocaine-Induced Chromatin Modifications Associate With Increased Expression and Three-Dimensional Looping of Auts2. *Biological Psychiatry*, *82*(11), 794–805. https://doi.org/10.1016/j.biopsych.2017.04.013

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

Feng, J., Shao, N., Szulwach, K. E., Vialou, V., Huynh, J., Zhong, C., … Nestler, E. J. (2015). Role of Tet1 and 5-hydroxymethylcytosine in cocaine action. *Nature Neuroscience*, *18*(4), 536–544. https://doi.org/10.1038/nn.3976

Feng, J., Wilkinson, M., Liu, X., Purushothaman, I., Ferguson, D., Vialou, V., … Shen, L. (2014). Chronic cocaine-regulated epigenomic changes in mouse nucleus accumbens. *Genome Biology*, *15*(4), R65. https://doi.org/10.1186/gb-2014-15-4-r65

Freeman, W. M., Lull, M. E., Patel, K. M., Brucklacher, R. M., Morgan, D., Roberts, D. C. S., & Vrana, K. E. (2010). Gene expression changes in the medial prefrontal cortex and nucleus accumbens following abstinence from cocaine self-administration. *BMC Neuroscience*, *11*, 29. https://doi.org/10.1186/1471-2202-11-29

Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., … Geschwind, D. H. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*, *362*(6420), eaat8127. https://doi.org/10.1126/science.aat8127

Garavan, H., Pankiewicz, J., Bloom, A., Cho, J.-K., Sperry, L., Ross, T. J., … Stein, E. A. (2000). Cue-Induced Cocaine Craving: Neuroanatomical Specificity for Drug Users and Drug Stimuli. *American Journal of Psychiatry*, *157*(11), 1789–1798. https://doi.org/10.1176/appi.ajp.157.11.1789

Gomez-Velazquez, M., Badia-Careaga, C., Lechuga-Vieco, A. V., Nieto-Arellano, R., Tena, J. J., Rollan, I., …

Manzanares, M. (2017). CTCF counter-regulates cardiomyocyte development and maturation programs in the embryonic heart. *PLOS Genetics*, *13*(8), e1006985. https://doi.org/10.1371/journal.pgen.1006985

Gross, J. A., Fiori, L. M., Labonté, B., Lopez, J. P., & Turecki, G. (2013). Effects of promoter methylation on increased expression of polyamine biosynthetic genes in suicide. *Journal of Psychiatric Research*, *47*(4), 513–519. https://doi.org/10.1016/j.jpsychires.2012.11.016

GTEx Consortium, T. Gte. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (New York, N.Y.)*, *348*(6235), 648–660. https://doi.org/10.1126/science.1262110

Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A., & Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols*, *6*(4), 468–481. https://doi.org/10.1038/nprot.2010.190

Guo, J. U., Su, Y., Shin, J. J. H., Shin, J. J. H., Li, H., Xie, B., … Song, H. (2013). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature Neuroscience*, *17*(2), 215–222. https://doi.org/10.1038/nn.3607

Hannon, E., Marzi, S. J., Schalkwyk, L. S., & Mill, J. (2019). Genetic risk variants for brain disorders are enriched in cortical H3K27ac domains. *Molecular Brain*, *12*(1), 7. https://doi.org/10.1186/s13041-019-0429-4

Hashimoto, H., Wang, D., Horton, J. R., Zhang, X., Corces, V. G., & Cheng, X. (2017). Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Molecular Cell*, *66*(5), 711-720.e3. https://doi.org/10.1016/j.molcel.2017.05.004

Iwata, A., Nagata, K., Hatsuta, H., Takuma, H., Bundo, M., Iwamoto, K., … Tsuji, S. (2014). Altered CpG methylation in sporadic Alzheimer's disease is associated with APP and MAPT dysregulation. *Human Molecular Genetics*, *23*(3), 648–656. https://doi.org/10.1093/hmg/ddt451

Johnson, A. R., Thibeault, K. C., Lopez, A. J., Peck, E. G., Sands, L. P., Sanders, C. M., … Calipari, E. S. (2019). Cues play a critical role in estrous cycle-dependent enhancement of cocaine reinforcement. *Neuropsychopharmacology*. https://doi.org/10.1038/s41386-019-0320-0

Kasper, C., Hebert, F. O., Aubin-Horth, N., & Taborsky, B. (2018). Divergent brain gene expression profiles between alternative behavioural helper types in a cooperative breeder. *Molecular Ecology*, *27*(20), 4136–4151. https://doi.org/10.1111/mec.14837

Kelz, M. B., Chen, J., Carlezon, W. A., Whisler, K., Gilden, L., Beckmann, A. M., … Nestler, E. J. (1999). Expression of the transcription factor ΔFosB in the brain controls sensitivity to cocaine. *Nature*, *401*(6750), 272–276. https://doi.org/10.1038/45790

Kim, T. H., Abdullaev, Z. K., Smith, A. D., Ching, K. A., Loukinov, D. I., Green, R. D., … Ren, B. (2007). Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. *Cell*, *128*(6), 1231–1245. https://doi.org/10.1016/j.cell.2006.12.048

Kozlenkov, A., Roussos, P., Timashpolsky, A., Barbu, M., Rudchenko, S., Bibikova, M., … Dracheva, S. (2014). Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. *Nucleic Acids Research*, *42*(1), 109–127.

https://doi.org/10.1093/nar/gkt838

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., … Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*(7539), 317–330. https://doi.org/10.1038/nature14248

LaPlant, Q., Vialou, V., Covington, H. E., Dumitriu, D., Feng, J., Warren, B. L., … Nestler, E. J. (2010). Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. *Nature Neuroscience*, *13*(9), 1137–1143. https://doi.org/10.1038/nn.2619

Li, X., Zhao, Q., Wei, W., Lin, Q., Magnan, C., Emami, M. R., … Bredy, T. W. (2019). The DNA modification N6-methyl-2'-deoxyadenosine (m6dA) drives activity-induced gene expression and is required for fear extinction. *Nature Neuroscience*, *22*(4), 534–544. https://doi.org/10.1038/s41593-019-0339-x

Lobo, M. K., & Nestler, E. J. (2011, July 18). The striatal balancing act in drug addiction: Distinct roles of direct and indirect pathway medium spiny neurons. *Frontiers in Neuroanatomy*. Frontiers Media SA. https://doi.org/10.3389/fnana.2011.00041

Luo, C., Keown, C. L., Kurihara, L., Zhou, J., He, Y., Li, J., … Ecker, J. R. (2017). Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science*, *357*(6351), 600–604. https://doi.org/10.1126/science.aan3351

Lutz, P. E., Tanti, A., Gasecka, A., Barnett-Burns, S., Kim, J. J., Zhou, Y., … Turecki, G. (2017). Association of a history of child abuse with impaired myelination in the anterior cingulate cortex: Convergent epigenetic, transcriptional, and morphological evidence. *American Journal of Psychiatry*, *174*(12), 1185–1194. https://doi.org/10.1176/appi.ajp.2017.16111286

Mash, D., Ffrench-Mullen, J., Adi, N., Qin, Y., Buck, A., Pablo, J., & ffrench-Mullen, Jarlath. (2007). Gene expression in human hippocampus from cocaine abusers identifies genes which regulate extracellular matrix remodeling. *{PloS} One*, *2*(11). https://doi.org/10.1371/journal.pone.0001187

Matsumoto, K., Nishihara, S., Kamimura, M., Shiraishi, T., Otoguro, T., Uehara, M., … Ogura, T. (2004). The prepattern transcription factor Irx2, a target of the FGF8/MAP kinase cascade, is involved in cerebellum formation. *Nature Neuroscience*, *7*(6), 605–612. https://doi.org/10.1038/nn1249

Maunakea, A. K., Chepelev, I., Cui, K., & Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Research*, *23*(11), 1256–1269. https://doi.org/10.1038/cr.2013.110

Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., … Costello, J. F. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, *466*(7303), 253–257. https://doi.org/10.1038/nature09165

Maurano, M. T., Wang, H., John, S., Shafer, A., Canfield, T., Lee, K., & Stamatoyannopoulos, J. A. (2015). Role of DNA Methylation in Modulating Transcription Factor Occupancy. *Cell Reports*, *12*(7), 1184–1195. https://doi.org/10.1016/j.celrep.2015.07.024

Nestler, E. J. (2005). The neurobiology of cocaine addiction. *Science & Practice Perspectives*, *3*(1), 4–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18552739

Nestler, E. J. (2014). Epigenetic mechanisms of drug addiction. *Neuropharmacology*, *76 Pt B*, 259–268.

https://doi.org/10.1016/j.neuropharm.2013.04.004

Rizzardi, L. F., Hickey, P. F., Rodriguez DiBlasi, V., Tryggvadóttir, R., Callahan, C. M., Idrizi, A., … Feinberg, A. P. (2019). Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability. *Nature Neuroscience*, *22*(2), 307–316. https://doi.org/10.1038/s41593-018-0297-8

Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G., & Ma'ayan, A. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database*, *2016*, baw100. https://doi.org/10.1093/database/baw100

Saftig, P., & Lichtenthaler, S. F. (2015). The alpha secretase ADAM10: A metalloprotease with multiple functions in the brain. *Progress in Neurobiology*, *135*, 1–20. https://doi.org/10.1016/j.pneurobio.2015.10.003

Sheffield, N. C., & Bock, C. (2016). LOLA: Enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics*, *32*(4), 587–589. https://doi.org/10.1093/bioinformatics/btv612

Shukla, M., Maitra, S., Hernandez, J.-F., Govitrapong, P., & Vincent, B. (2019). Methamphetamine regulates βAPP processing in human neuroblastoma cells. *Neuroscience Letters*, *701*, 20–25. https://doi.org/10.1016/j.neulet.2019.02.023

Tena, J. J., Alonso, M. E., de la Calle-Mustienes, E., Splinter, E., de Laat, W., Manzanares, M., & Gómez-Skarmeta, J. L. (2011). An evolutionarily conserved three-dimensional structure in the vertebrate Irx clusters facilitates enhancer sharing and coregulation. *Nature Communications*, *2*(1), 310. https://doi.org/10.1038/ncomms1301

Tepper, J. M., Tecuapetla, F., Koós, T., & Ibáñez-Sandoval, O. (2010). Heterogeneity and diversity of striatal GABAergic interneurons. *Frontiers in Neuroanatomy*, *4*(December), 150. https://doi.org/10.3389/fnana.2010.00150

Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., … Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research*, *13*(9), 2129–2141. https://doi.org/10.1101/gr.772403

Vaillancourt, K., Ernst, C., Mash, D., & Turecki, G. (2017). DNA Methylation Dynamics and Cocaine in the Brain: Progress and Prospects. *Genes*, *8*(5), 138. https://doi.org/10.3390/genes8050138

Vojta, A., Dobrinić, P., Tadić, V., Bočkor, L., Korać, P., Julg, B., … Zoldoš, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Research*, *44*(12), 5615–5628. https://doi.org/10.1093/nar/gkw159

Volkow, N. D., Wang, G.-J., Telang, F., Fowler, J. S., Logan, J., Childress, A.-R., … Wong, C. (2006). Cocaine Cues and Dopamine in Dorsal Striatum: Mechanism of Craving in Cocaine Addiction. *Journal of Neuroscience*, *26*(24), 6583–6588. https://doi.org/10.1523/JNEUROSCI.1544-06.2006

Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T. W., Greven, M. C., … Weng, Z. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Research*, *22*(9), 1798–1812. https://doi.org/10.1101/GR.139105.112

Yang, X., Han, H., De Carvalho, D. D., Lay, F. D., Jones, P. A., & Liang, G. (2014). Gene Body Methylation Can Alter Gene Expression and Is a Therapeutic Target in Cancer. *Cancer Cell*, *26*(4), 577–590. https://doi.org/10.1016/j.ccr.2014.07.028

Yao, P., Lin, P., Gokoolparsadh, A., Assareh, A., Thang, M. W. C., & Voineagu, I. (2015). Coexpression networks identify brain region–specific enhancer RNAs in the human brain. *Nature Neuroscience*, *18*(8), 1168–1174. https://doi.org/10.1038/nn.4063

Zhang, T. Y., Hellstrom, I. C., Bagot, R. C., Wen, X., Diorio, J., & Meaney, M. J. (2010). Maternal care and DNA methylation of a glutamic acid decarboxylase 1 promoter in rat hippocampus. *Journal of Neuroscience*, *30*(39), 13130–13137. https://doi.org/10.1523/JNEUROSCI.1039-10.2010

Zhang, T. Y., Keown, C. L., Wen, X., Li, J., Vousden, D. A., Anacker, C., … Meaney, M. J. (2018). Environmental enrichment increases transcriptional and epigenetic differentiation between mouse dorsal and ventral dentate gyrus. *Nature Communications*, *9*(1), 1–11. https://doi.org/10.1038/s41467-017-02748-x

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 of the United States of America*, *108*(16), 6626–6631. https://doi.org/10.1073/pnas.1018514108

7. Figures

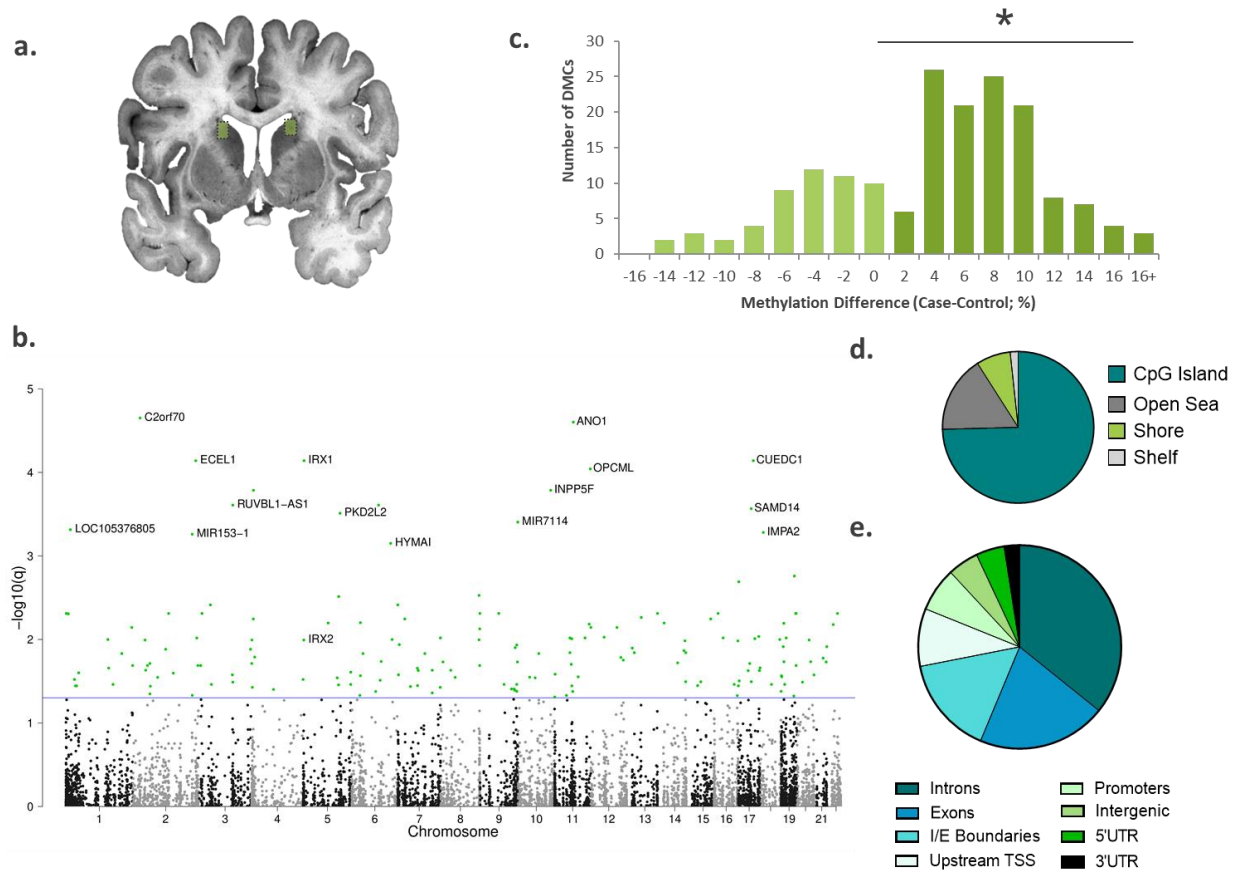


Figure 1. Widespread changes in DNA methylation associated with chronic cocaine dependence in the human caudate nucleus. a) Dissections from dorsolateral caudate nucleus were used for reduced representation bisulfite sequencing (n=25 per group, boxed). b) Manhattan plot showing the chromosomal location of all significantly differentially methylated CpG regions (DMRs); blue line represents FDR $q < 0.05$. c) Although hyper- and hypomethylated DMCs were identified, there was a significant bias towards clusters with increased methylation in the cocaine group $\chi^2 = 26.575$; * $p < 0.05$. d)Most of the DMCs overlap with known CpG islands and e) annotated introns, exons and intron-exon boundaries.

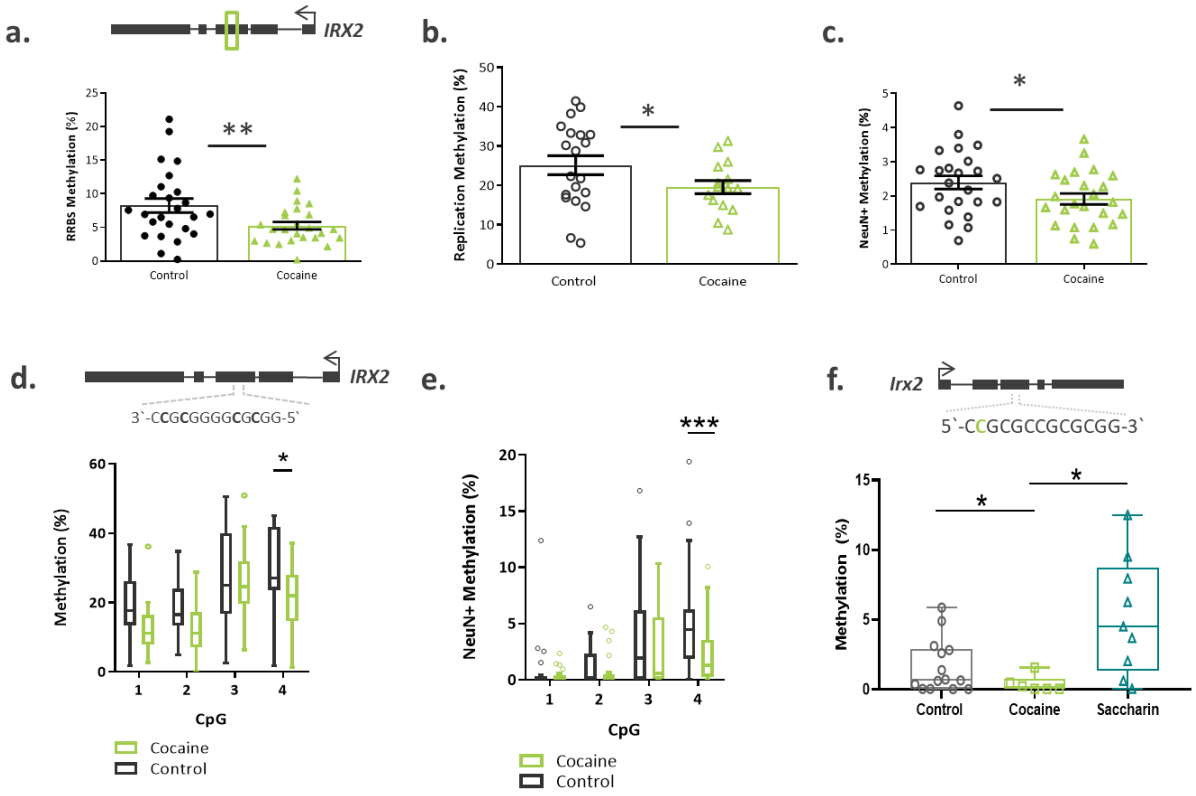


Figure 2. *IRX2* is hypomethylated in the caudate nucleus. a) RRBS analysis identified a cluster of 21 CpGs within the third exon of *IRX2* that were less methylated in the cocaine group of the discovery cohort (n=25 per group). b) This data was replicated in an independent cohort of caudate samples (n=15 cases and n=20 controls). c) Hypomethylation was specific to neuronal (NeuN+) nuclei (n=24 cases and n=24 controls; discovery cohort). d) The 5' CpG within the CTCF binding site (exon 3) was hypomethylated in the cocaine group in caudate tissue homogenate (n= 20 controls and 17 cases, replication cohort) and e) neuronal nuclei (n=24 controls and n=23 cases). f) The 5' most CpG in the mouse CTCF site was significantly less methylated after cocaine self-administration (n=6), compared to non-drug reward self-administration (n=9) or controls (n=14). Box plots indicate mean and range of data. Bar data represented as mean ± s.e.m. * $p < 0.05$; ** q val < 0.02

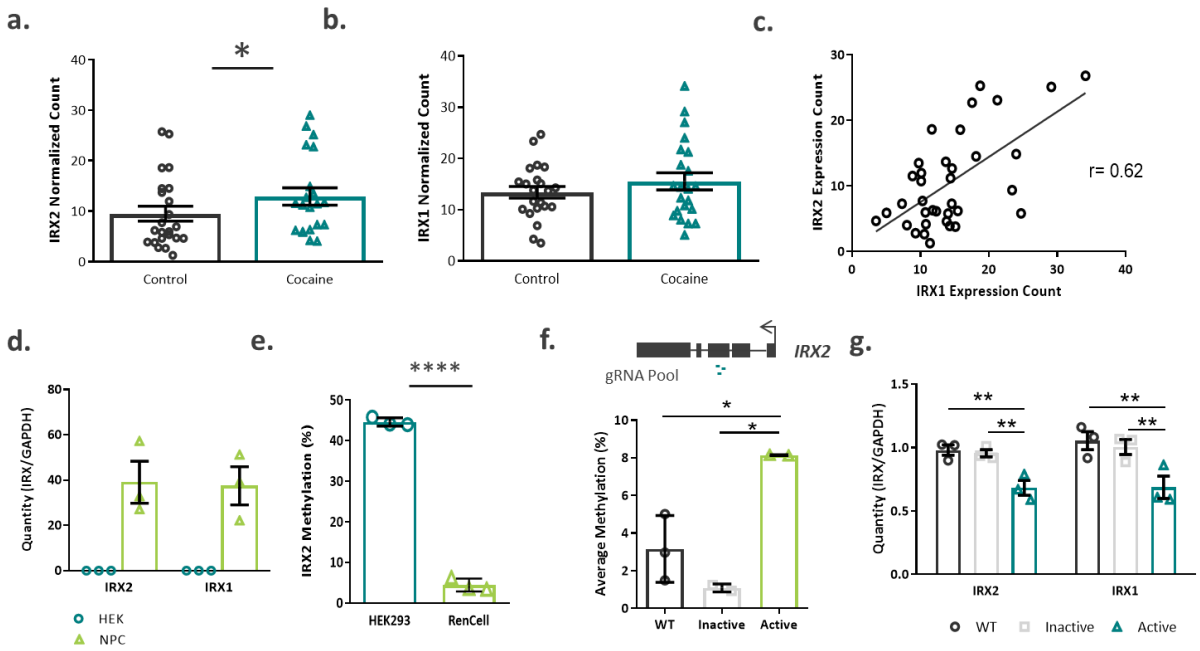


Figure 3. *IRX2* expression is increased in cocaine use disorder and is related to exon 3 methylation in cells. a) *IRX2* expression was significantly increased in the caudate nucleus of cocaine dependent subjects (n=21 cases and n=23 controls). b) while no significant increase in *IRX1* was detected (n=22 cases and n=22 controls), c) the expression of both *IRX1* and *IRX2* transcripts was highly correlated (n=36). d) Human neural progenitor cells (RENcells) endogenously expressed *IRX1* and *IRX2* while kidney epithelial cells (HEK293) did not (n=3 per group). e) Endogenous methylation of *IRX2* exon 3 was higher in kidney epithelial cells than in neural progenitor cells (n=3 per group). f) Transfection of an active dCas9-DNMT3A construct, along with a pool of 3 guide RNA constructs significantly increased methylation of *IRX2* exon 3 in RenCells, compared to transfection with an inactive construct or wildtype controls. g) Active methylation of *IRX2* decreased transcription of both *IRX2* and *IRX1* compared to inactive or wildtype cells. (n=2-3 replicates per group) WT=wildtype. Data represented as mean ± s.e.m **** p <0.0001, * p < 0.05

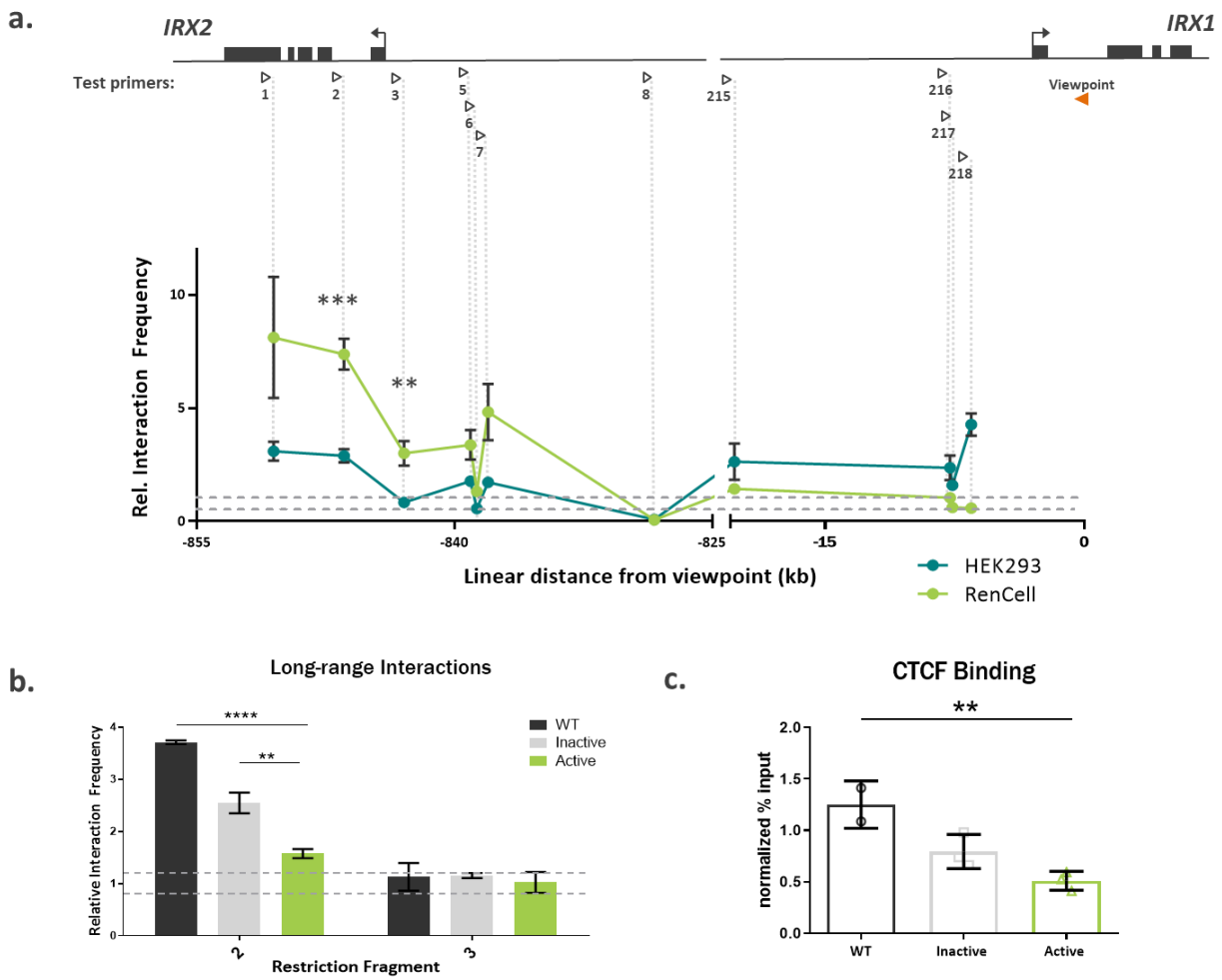


Figure 4. Long range chromatin structure of the *IRXA* gene cluster is impacted by methylation. a) Chromatin conformation capture (3C) experimental design contained a standard viewpoint within the first intron of *IRX1* (orange arrowhead), and test primers (open arrowheads) tiled across IRX2 and the intergenic region between the genes. The genomic fragments containing *IRX2* functionally interacted with the 5' end *IRX1* in two human cell types, with significantly higher rates of interaction observed in RenCells. The noise band, where interaction frequencies would be expected by chance ("random collisions"), is indicated by the horizontal dashed lines. Vertical dashed lines pair restriction fragment with its corresponding data points. N=3 replicates per group. b) Active methylation of HEK293 cells decreased long range interaction between restriction fragment 2 and the *IRX1* gene (n=3 replicates of 10 x 10^6 cells per group). c) The dCas9-DNMT3A transfection significantly (n=3 replicates of 5 x 10^6 cells per group) decreased CTCF binding to IRX2 exon 3 compared to wildtype cells. WT=wildtype. Data represented as mean ± s.e.m. ** $p < 0.01$; *** $p < 0.0005$

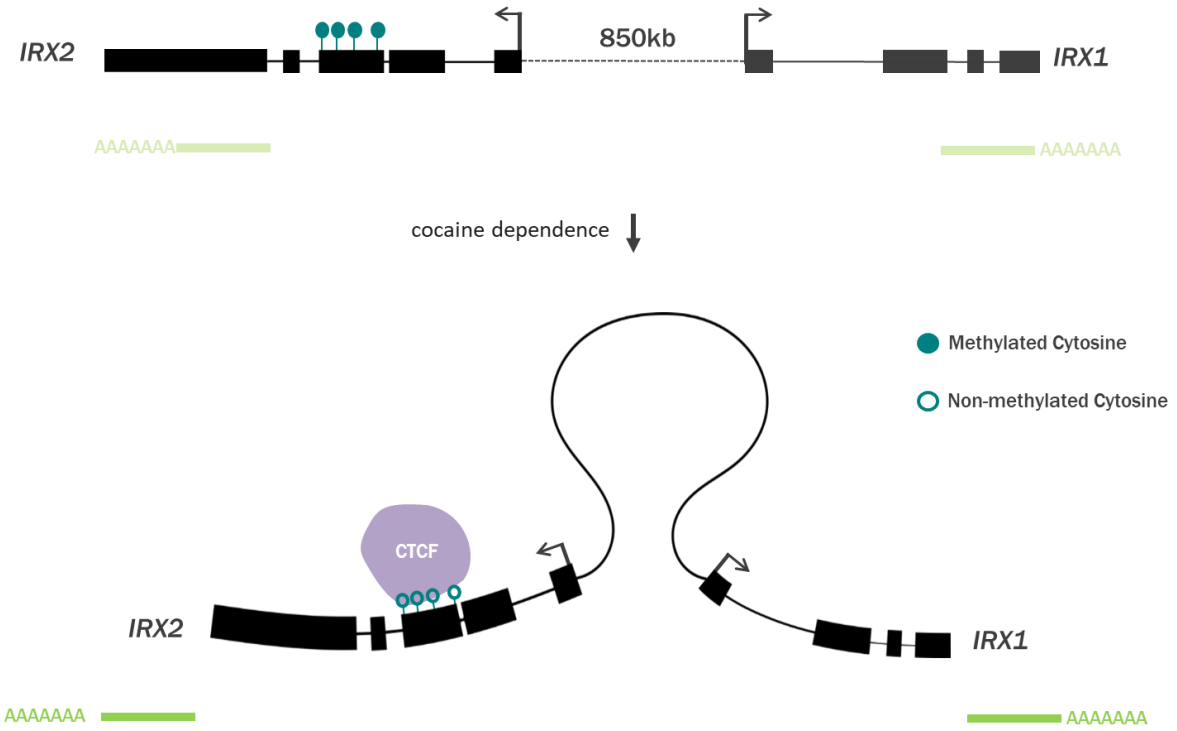


Figure 5. **A model for cocaine-sensitivity of 3D chromatin organization at the *IRXA* gene cluster.** Cocaine dependence is associated with decreased intragenic methylation of *IRX2*, which may increase *IRXA* gene expression through CTCF-mediated chromatin architecture.

8. Tables

Table 1. Differentially methylated regions with nominally differential gene expression

Chr	From	To	Gene	Number of CpGs	Methylation Difference	RNAseq Fold Change	RNAseq p-value
5	2748781	2748955	IRX2	21	-3.06%	1.20	8.17E-02
9	126776177	126776282	LHX2	12	-8.22%	1.22	1.46E-02
5	3599609	3599704	IRX1	9	11.05%	1.39	1.21E-03
4	53474	53566	ZNF595	9	7.67%	0.79	1.08E-02
17	1960987	1961029	HIC1	8	-11.51%	1.25	7.09E-03
2	26785211	26785290	C2orf70	7	7.44%	1.29	1.31E-02
1	17215449	17215492	CROCC	7	-5.45%	1.24	1.85E-03
11	132812684	132812729	OPCML	7	15.93%	0.86	7.03E-02
2	63274825	63274901	OTX1	6	9.92%	1.25	2.53E-02
18	77918229	77918253	PARD6G	6	-10.83%	1.51	3.98E-05
5	131607235	131607278	PDLIM4	6	8.28%	1.36	2.78E-03
7	150002	150037	AC093627.10	5	8.95%	0.84	2.14E-02
12	120654707	120654747	PXN	5	10.21%	1.17	5.55E-02
19	18980163	18980188	UPF1	5	10.03%	1.07	2.35E-02
2	63275003	63275040	OTX1	4	7.49%	1.25	2.53E-02
17	76172805	76172850	TK1	4	7.10%	0.87	8.77E-02
8	37556087	37556121	ZNF703	4	12.14%	1.20	1.41E-02
6	32165134	32165176	NOTCH4	3	8.58%	1.14	9.81E-02
3	8799985	8800008	OXTR	3	12.55%	1.29	1.35E-02
19	47220817	47220856	PRKD2	3	5.76%	1.33	3.70E-04
5	176877611	176877636	PRR7	3	11.64%	1.53	4.52E-11
4	6273547	6273577	WFS1	3	21.42%	1.12	7.12E-02
9	136654410	136654426	VAV2	2	12.46%	1.17	3.08E-02

9. Supplemental Tables and Figures

Supplemental Table 1. Demographic Information for Discovery Cohort

	Controls	Cases	Effect Size***
N	25	25	
Ethnicity*	11 (A); 7 (C); 11 (H)	13 (A); 6 (C); 10 (H)	
Age (years)	38.4 ± 12.5	35.2 ± 8.5	Small (0.297)
PMI (hours)	15.1 ± 7.2	14.5 ± 7.7	Small (0.117)
pH	6.4 ± 0.3	6.5 ± 0.3	Small (0.128)
Sex**	25 (M)	25 (M)	

*A= African American; C= Caucasian; H= Hispanic

** M=male

*** Calculated with Cohen's d

Supplemental Table 2. Sequencing Statistics from RRBS experiment

Group	Bisulfite Conversion Efficiency (% mean ± s.d)	Average Reads per Library	Alignment (%)	CpG Coverage at 1X (#)	CpG Coverage at 5X (#)	CpG Coverage at 10X (#)
Cocaine	98.99 ± 0.215	2.59E+07	59.6	2.19E+06	1.50E+06	1.12E+06
Control	99.04 ±0.199	3.58E+07	59.2	2.29E+06	1.63E+06	1.27E+06

Supplemental Table 3. Differentially Methylated Regions from RRBS experiment

Chromosome	From	To	Closest Gene	Methylation Difference
1	1957219	1957242	GABRD	-13.56
1	9384997	9385034	SPSB1	11.42
1	17215449	17215492	LOC105376805	-6.53
1	32052631	32052670	TINAGL1	9.19
1	36184567	36184613	C1orf216	-1.09
1	40105638	40105680	HEYL	7.62
1	47899819	47899864	FOXD2	7.20
1	156370464	156370511	C1orf61	-3.59
1	159158440	159158481	CADM3	3.53
1	175474534	175474573	TNR	-9.97
1	207842941	207842986	CR1L	2.06
1	245081971	245082017	-	2.11
1	248100364	248100422	OR2L13	6.59
2	26785211	26785290	C2orf70	6.30
2	45179874	45179917	-	6.26
2	47597161	47597192	EPCAM	-10.81
2	54086813	54086908	PSME4	1.14
2	63274825	63274901	OTX1	7.26
2	63275003	63275040	OTX1	7.93
2	64978656	64978784	-	1.47
2	91776876	91777107	-	4.50
2	121584523	121584544	GLI2	10.68
2	132559342	132559383	C2orf27B	-5.33

2	150177031	150177122	-	6.69
2	220159870	220159964	MIR153-1	3.89
2	220313302	220313338	SPEG	8.91
2	233352680	233352722	ECEL1	2.83
2	237278447	237278472	IQCA1	8.56
2	241190895	241190939	-	9.35
3	8799985	8800008	OXTR	12.09
3	13261073	13261093	-	17.44
3	44622590	44622600	ZNF660	-6.69
3	48646667	48646734	SNORA94	-1.85
3	126075000	126075041	KLF15	7.72
3	127795875	127795920	RUVBL1-AS1	4.56
3	128035446	128035489	EEFSEC	15.09
3	193587448	193587494	-	6.68
4	53474	53566	ZNF718	7.39
4	1027718	1027766	-	8.38
4	2366616	2366653	ZFYVE28	5.87
4	6273547	6273577	WFS1	22.09
4	6660521	6660586	-	-4.94
4	11370466	11370514	MIR572	5.92
4	81119105	81119153	PRDM8	2.59
4	190935969	190936172	-	5.08
5	2748781	2748955	IRX2	-2.78
5	3599609	3599704	IRX1	10.68
5	92908052	92908070	NR2F1-AS1	5.85

5	126565076	126565118	-	1.09
5	131607235	131607278	PDLIM4	8.22
5	132159035	132159058	SHROOM1	9.71
5	137225015	137225062	PKD2L2	-2.84
5	176794630	176794733	RGS14	-0.54
5	176877611	176877636	PRR7	11.09
6	30095099	30095213	-	5.55
6	32118204	32118299	PRRT1	-1.40
6	32165134	32165176	NOTCH4	7.99
6	41516138	41516182	FOXP4	13.36
6	88757352	88757398	SPACA1	-6.14
6	99291702	99291749	-	2.19
6	105389236	105389283	LINC00577	-1.03
6	109038319	109038353	-	11.26
6	144329386	144329433	HYMAI	-6.32
6	149772252	149772346	ZC3H12D	2.29
7	150002	150037	LOC100507642	8.38
7	2681444	2681481	TTYH3	-6.28
7	5535370	5535410	MIR589	5.32
7	25892334	25892381	-	13.57
7	47092583	47092674	-	5.33
7	72838672	72838715	-	6.05
7	73457700	73457743	ELN	-7.80
7	75911977	75912017	SRRM3	3.02
7	100463903	100463923	TRIP6	-5.91

7	127991684	127991788	PRRT4	-1.72
7	128045595	128045723	IMPDH1	-3.98
7	157484891	157484931	PTPRN2	-3.61
7	158030795	158030841	PTPRN2	13.93
8	8559811	8559854	CLDN23	-5.22
8	37556087	37556121	ZNF703	13.16
8	53854332	53854406	NPBWR1	1.00
8	142386495	142386530	-	-12.01
8	142411559	142411588	-	11.10
8	145018229	145018256	MIR661	7.14
8	145047853	145047897	PARP10	-0.03
8	145557189	145557231	SCRT1	4.45
9	69785442	69785487	-	7.93
9	89260398	89260444	-	9.52
9	116930259	116930276	COL27A1	6.16
9	126776177	126776282	LHX2	-6.50
9	128998377	128998421	-	-13.96
9	130517323	130517347	SH2D3C	5.99
9	136294824	136294869	ADAMTS13	1.99
9	136437916	136437929	ADAMTSL2	6.50
9	136444285	136444322	FAM163B	8.52
9	138881845	138881947	-	3.44
9	140343322	140343368	MIR7114	-14.08
10	43361570	43361618	-	2.15
10	63809073	63809121	ARID5B	2.33

10	77191186	77191233	C10orf11	4.17
10	121578031	121578079	INPP5F	7.31
11	830237	830321	CRACR2B	2.50
11	2188129	2188165	TH	9.09
11	44325704	44325748	ALX4	4.84
11	57194509	57194551	SLC43A3	7.81
11	57243954	57244075	RTN4RL2	2.60
11	62370210	62370242	EML3	3.74
11	64068355	64068401	TEX40	4.82
11	64405907	64405955	NRXN2	-9.25
11	67778106	67778151	ALDH3B1	5.81
11	69924989	69925068	ANO1	2.60
11	87908343	87908437	MIR3166	-3.53
11	113185495	113185541	TTC12	-3.61
11	128694133	128694170	-	2.72
11	132812684	132812729	OPCML	14.22
12	312616	312661	LOC101929384	5.78
12	49373486	49373526	WNT1	6.85
12	97379994	97380030	-	14.41
12	111856097	111856140	SH2B3	9.46
12	120654707	120654747	PXN	11.34
13	20751530	20751578	-	5.85
13	28369133	28369177	GSX1	9.07
13	54357765	54357789	-	11.42
13	114516850	114516880	GAS6-AS1	7.76

14	24045352	24045400	JPH4	7.42
14	74252858	74252903	LOC100506476	-7.07
14	97499801	97499847	-	1.45
14	101370936	101370984	MEG8	-3.58
14	104182037	104182078	ZFYVE21	2.87
14	104561244	104561286	ASPG	6.29
15	41218352	41218393	DLL4	2.57
15	53077859	53077900	ONECUT1	4.44
15	89921805	89921837	MIR9-3HG	9.22
16	15083748	15083904	PDXDC1	3.07
16	33961382	33961409	LINC00273	8.67
16	75148495	75148543	LDHD	-6.65
16	86528713	86528754	FENDRR	8.75
16	88677757	88677805	ZC3H18	7.63
17	259900	259942	C17orf97	2.77
17	1960987	1961029	HIC1	-9.42
17	3795770	3795869	P2RX1	-0.76
17	36734981	36735041	SRCIN1	-4.37
17	42431309	42431351	FAM171A2	-5.54
17	48194729	48194774	SAMD14	5.08
17	48351738	48351753	TMEM92	3.74
17	55939256	55939304	CUEDC1	-2.65
17	61524397	61524469	CYB561	8.79
17	76172805	76172850	TK1	3.01
17	78485795	78485834	-	14.00

17	81039201	81039245	METRNL	-5.35
18	11999676	11999716	IMPA2	21.97
18	42324761	42324793	SETBP1	12.05
18	72916808	72916831	ZADH2	-4.46
18	77918229	77918253	PARD6G	-10.96
19	1295710	1295735	EFNA2	-9.55
19	2252030	2252070	JSRP1	6.92
19	7934653	7934697	PRR36	-1.48
19	10023964	10024010	OLFM2	-4.29
19	10543010	10543120	PDE4A	4.45
19	13264508	13264657	IER2	2.32
19	18980163	18980188	GDF1	8.39
19	47220817	47220856	STRN4	9.06
19	49340489	49340532	PLEKHA4	-0.89
19	54057910	54057958	ZNF331	-5.53
19	55598698	55598739	PPP1R12C	-5.60
19	55685105	55685146	SYT5	3.59
20	20433253	20433298	RALGAPA2	-5.18
20	44879878	44879903	CDH22	9.00
20	50721767	50721789	ZFP64	-3.58
20	62948105	62948151	-	-2.71
21	33941737	33941824	-	8.09
21	45232224	45232267	AATBC	-14.09
21	47580847	47580889	SPATC1L	14.06
22	23441360	23441403	GNAZ	4.89

22	39096951	39096999	GTPBP1	2.70

Supplemental Table 4. Demographic Information for Replication Cohort

	Controls	Cases
N	21	15
Ethnicity*	21 (C)	15 (C)
Age (years)	47.3 ± 21.9	41.7 ± 9.9
PMI (hours)	38.7 ± 25.3	53.3 ± 24.1
pH	6.5 ± 0.2	6.4 ± 0.4
Sex**	20 (M); 1 (F)	13 (M); 2(F)
Co-dependence***	0	6 (A); 2(C)

* C= Caucasian

** M = Male, F= Female

*** A=Alcohol ; C=Cannabis

Supplemental Table 5. Probes used for nanoString experiment

Gene	Probe Sequence
IRX1	5'-AGGGAGGGAATGTGGGAGGAATTAAGACAAATATTTCAGACTGGTGTAAAGGACAAATATGACAACGACGTCAAGGACTCGCATCCGTCGCTTTCTGCAG-3'
IRX2	5'-ATGTCTAAATCCAGACCACATCTTGTGCCACTGTAAAAGAGAAGGATCCACACAGCACTGTCAGCCCACAGACTCTAACAATTGGAACAGACTATTGTTG-3'

Supplemental Table 6. Bisulfite amplicon sequencing primers

Primer	Direction	Primer Sequence	Predicted Length
Human (hg19_chr5:2748634-2749159)			
IRX2_1	F	5'-TAG GTT YGG TTG TTT GAG GT-3'	365bp
	R	5'-TAA ACT CRC TCA CRA ATC ACT-3'	
IRX2_2	F	5'-TTY GTA GTT TGT GTA GTT GTY GT-3'	480bp
	R	5'-AAC RAA AAA AAA CTT CCR TAC C-3'	
IRX2_3	F	5'-GTA GGG TAG GGY GAG TTT TTT-3'	459bp
	R	5'-CCT ACA CRT AAA CTC RCT CA-3'	
Human (hg19_chr5:3599609-3599704)			
IRX1_1	F	5'-TTT GGT TTC GGG TAA GAG TAG A-3'	315bp
	R	5'-TGG TGA TGA TGG TTA GTA TGA-3'	
IRX1_2	F	5’ GGG TAA GAG TAG AAG GGG TTG-3’	372bp
	R	5'-TGT TCG TTG AGT TAG GTT TTG A-3'	
IRX1_3	F	5'-GTT AGG TTT AAG GGC GAG TTT-3'	373bp
	R	5'-TTG GTG ATG ATG GTT AGT ATG A-3'	
Mouse (mm10_chr13:72630968-72632147)			
Irx2_1	F	5'-GAT AAG TTT GAG GAT TTG GAG G-3'	371bp
	R	5'-CCC AAC AAA AAC RAA ACA A-3'	
Irx2_2	F	5'-TGT AAG GAT AAG TTT GAG GAT TTG-3'	277bp
	R	5'-CCA AAC TTA ATT ACT TAA AAT CCR A-3'	
Irx2_3	F	5'-TAA GTT TGA GGA TTT GGA GGA-3'	278bp
	R	5'-AAC CCR AAC CCA AAC TTA A-3'	

Supplemental Table 7. Bisulfite amplicon PCR conditions

Temperature (°C)	Time (min:sec)	
95	05:00	
95	00:15	x 30
60	00:30	
72	00:30	
72	02:00	
4	∞	

Supplemental Table 8. qPCR probe and primer information

IRX1

Forward	5`- CCA GGG TTG TCC TTC AGT T – 3`
Reverse	5`- CCC AAC TAC AGC GCC TTC – 3`
Probe	5`- /56-FAM/TGC GAG CCC ATC TGC GAG AA/36-TAMSp/ - 3`

IRX2

Forward	5`- GTA CGG GTG GTA GCT GAT G – 3`
Reverse	5`- AGC CCG CTG CAG TAC TC – 3`
Probe	5`- /56-FAM/TGG TGT GCG CGT CGT AGG G/36-TAMSp – 3`

Supplemental Table 9. Primers used in dCas9-DNMT3A epigenome editing

Name	Purpose	Sequence
IRX2-sgRNA1	Guide RNA	5'-GTGACCTCGTCGCCGCTTAC-3'
IRX2-sgRNA2	Guide RNA	5'-TCCTGGCGGTGCCCCGGTTG-3'
IRX2-sgRNA3	Guide RNA	5'-ACAGCTTGGGCTTGCTGGCG-3'
U6	Construct sequencing primer	5'-GAGGGCCTATTTCCCATGATTCC -3'
Cas9-F	Relative plasmid quantification	5'-TGCCCCAAGTGAATATCGTG-3'
Cas9-R	Relative plasmid quantification	5'-GACTTGCCCTTTTCCACTTTG-3'
RAG1-F	Relative plasmid quantification	5'-TGTTGACTCGATCCACCCCA-3'
RAG1-R	Relative plasmid quantification	5'-TGAGCTGCAAGTTTGGCTGAA-3'

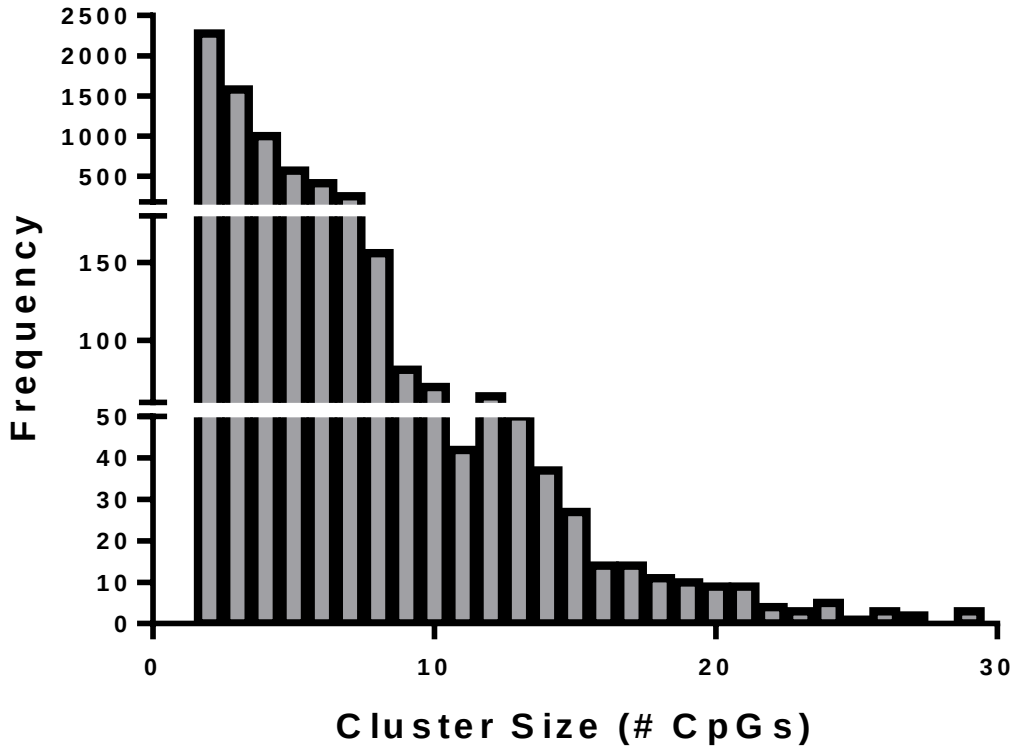
Supplemental Table 10. ChIP-qPCR Primers

Primer	Direction	Sequence	Amplicon Length
Upstream_Control	F	5'-CTACATGGGCGCACCCTAC-3'	103bp
	R	5'-GGGGTCGTTGAGCTGGTACG-3'	
Exon3BindingSite	F	5'-GCTTACCGGCTTGGAGGC-3'	77bp
	R	5'-CCTGGGGCGTCTTGCG-3'	

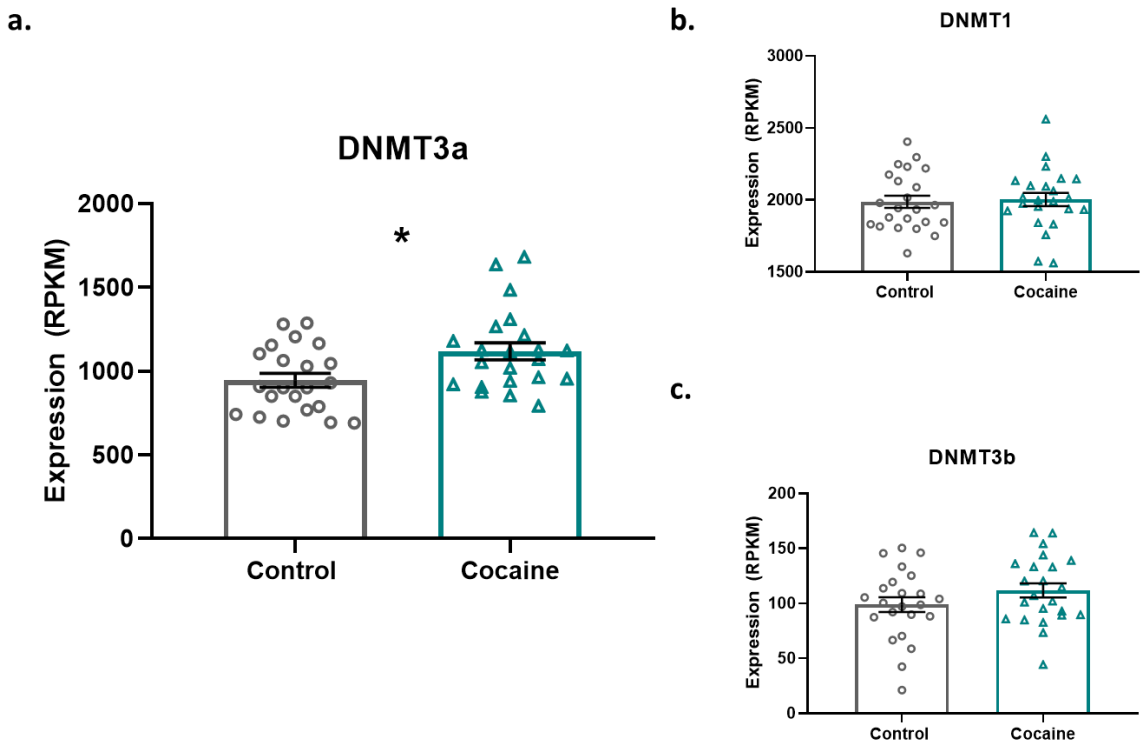
Supplemental Table 11. Primers used in 3C-qPCR experiment and quality control

Primer	Direction	Primer Sequence	Amplicon Length
3C-qPCR Primers			
IRX3C_probe		/56-FAM/AGA AGC AGG CCC GTG GAG CGG T/3TAMRA/	
IRX3C_C220	Reverse Anchor	5'-GACCTGCAGGAACCCCGTG-3'	
IRX3C_T1	Test Forward	5'-TCACTAGCGCTCTACCC-3'	184bp
IRX3C_T2	Test Forward	5'-CTGCTCTAAAGCTGGGCAAATC-3'	129bp
IRX3C_T3	Test Forward	5'-TCGACAGTCTGCGACATT-3'	111bp
IRX3C_T4	Test Forward	5'-ACGTGCTTCCTGAAGTTTCTCT-3'	121bp
IRX3C_T5	Test Forward	5'-GCCGCACATGTAAATAACAGAGA-3'	118bp
IRX3C_T6	Test Forward	5'-GACTGCAAGCCTCCTCCAC-3'	119bp
IRX3C_T7	Test Forward	5'-GCATGTCTTGGAATTGCCATTTA-3'	108bp
IRX3C_T8	Test Forward	5'-ACACCTGCTATAATTTGATTGGAA-3'	126bp
IRX3C_T215	Test Forward	5'-CTTTAAGCATAACCCATTTCGTCTTT-3'	177bp
IRX3C_T216	Test Forward	5'-GAGACATGCTACTATAACCATAATTGA-3'	186bp
IRX3C_T217	Test Forward	5'-TGACAATAGTTCATGAGCTTACAA-3'	145bp
IRX3C_T218	Test Forward	5'-ATGAGCTATCTCTTGCTCCTGA-3'	118bp
Internal Control Primers			
3C_GAPDH	F	5'-ACAGTCCATGCCATCACTGCC-3'	265bp
	R	5'-GCCTGCTTCACCACCTTCTTG-3'	
Digestion Efficiency Primers			
IRX_Dig1-2	F	5'-TCACTAGCGCTCTACCC-3'	149bp
	R	5'-GCTGGTTATACAATCTTTGTTGGT-3'	
IRX_Dig2-3	F	5'CTGCTCTAAAGCTGGGCAAATC-3'	126bp
	R	5'-GGCTGCAAATGCCTCGAATTTA-3'	
IRX_Dig3-4	F	5'-TCGACAGTCTGCGACATT-3'	128bp
	R	5'-GGAACCCTAAAGGTTGCCATC-3'	
IRX_Dig4-5	F	5'-ACGTGCTTCCTGAAGTTTCTCT-3'	73bp
	R	5'-CGGAGTCGGCGAGATGC-3'	
IRX_Dig5-6	F	5'-GCCGCACATGTAAATAACAGAGA-3'	189bp
	R	5'-TGTGAAGAGGCATCTGCTGT-3'	
IRX_Dig6-7	F	5'-GACTGCAAGCCTCCTCCAC-3'	130bp
	R	5'-GAGGCACTCAAGTTGGTACAAG-3'	
IRX_Dig7-8	F	5'-GCATGTCTTGGAATTGC-3'	153bp
	R	5'-TGAGCAAAGTGGGAACCAGG-3'	
IRX_Dig8-9	F	5'-ACACCTGCTATAATTTGATTGGAA-3'	179bp
	R	5'-ATGTGTTCAATATCTGCACACT-3'	
IRX_Dig215-216	F	5'-CTTTAAGCATAACCCATTTCGTCTTT-3'	167bp
	R	5'-TGCCGTTGGCACATATCCTC-3'	

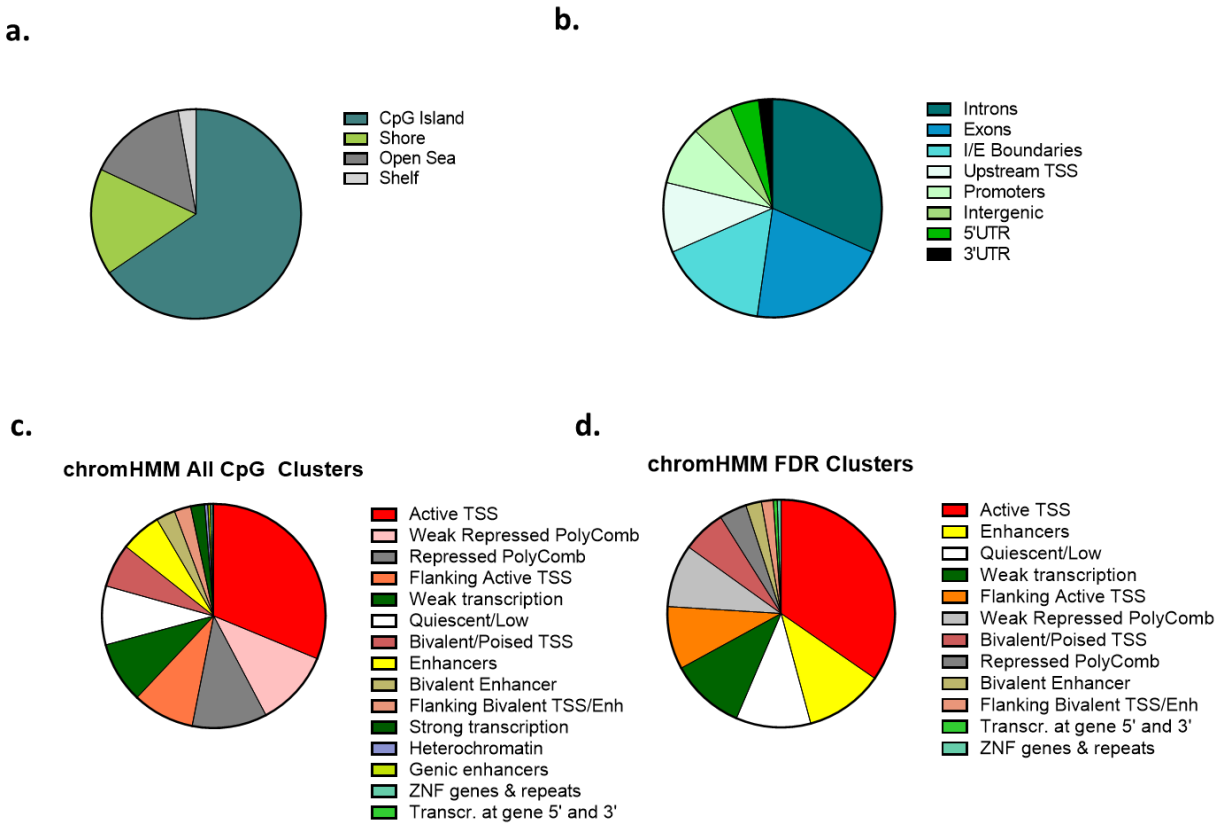
IRX_Dig216-217	F	5'-GAGACATGCTACTATAACCATAATTGA-3'	177bp
	R	5'-GGAAGGGGTTTAGAAAATGCAACA-3'	
IRX_Dig217-218	F	5'-TGACAATAGTTCATGAGCTTACAA-3'	145bp
	R	5'-GGCAGAGAGCATTTCACCCT-3'	
IRX_Dig218-219	F	5'-ATGAGCTATCTCTTGCTCCTGA-3'	141bp
	R	5'-GGAGACATCCCCTCTCCGT-3'	



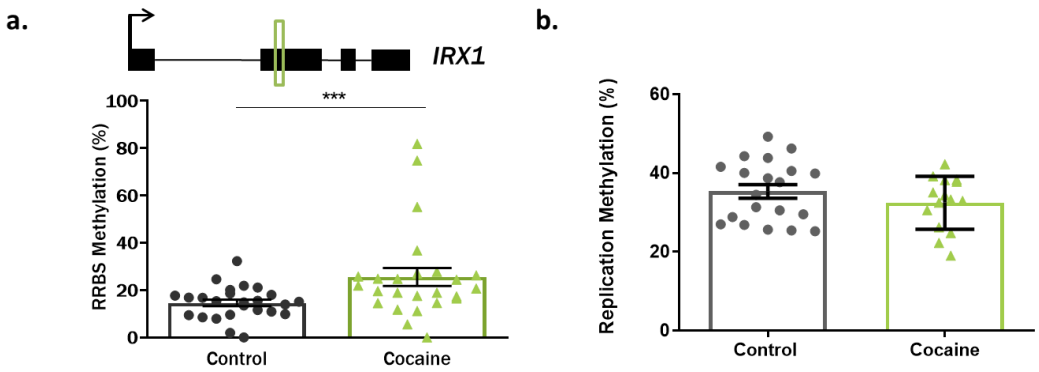
Supplemental Figure 1. Distribution of CpG cluster sizes.



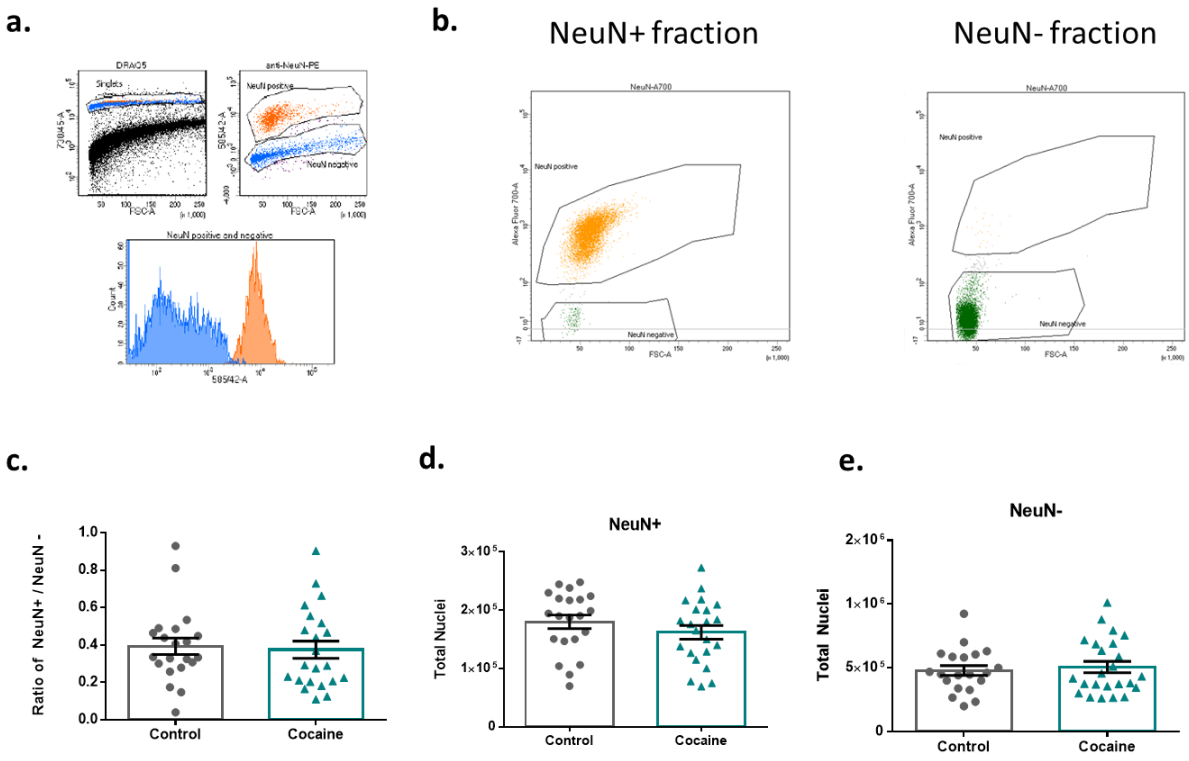
Supplemental Figure 2. Expression of DNA methyltransferase genes. a) The abundance of DNMT3a transcripts was higher in the cocaine group (n=22) compared to controls (n=22). b) There was no difference in expression of DNMT1 or c) DNMT3b (n=23 per group). Data represented as mean ± s.e.m. RPKM= reads per kilobase, per million reads. * $p < 0.05$



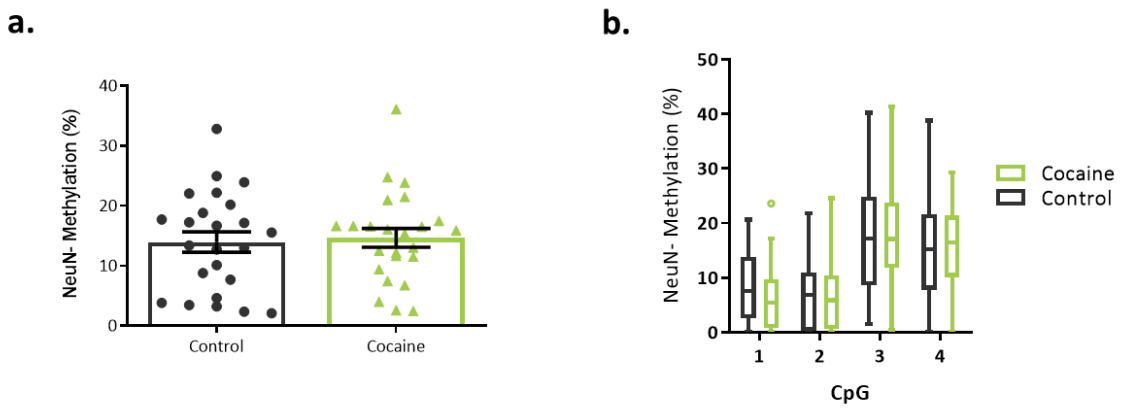
Supplemental Figure 3. The genomic and chromatin state contexts of all 6712 CpG clusters. Most detected CpG regions ranged in size from 2-10 CpGs and overlapped with b) The CpG clusters generated from RRBS overlap with a) known CpG islands and b) introns, exons and intron-exon boundaries. c) chromHMM chromatin state annotations for all identified clusters and d) clusters that were differentially methylated between cases and controls.



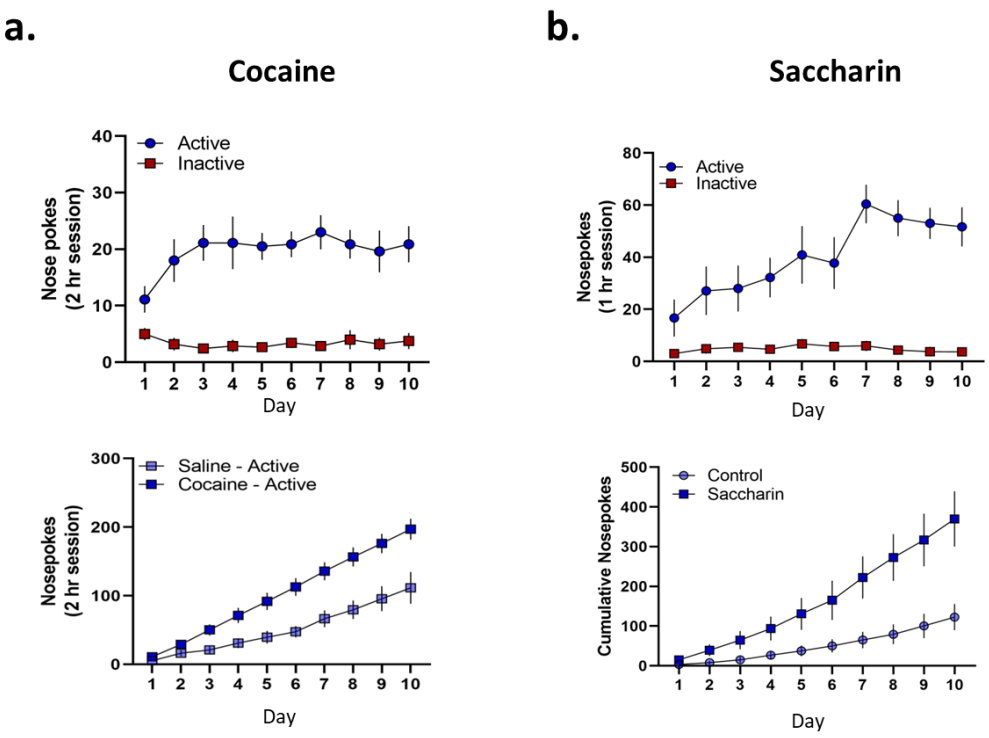
Supplemental Figure 4. *IRX1* exon 2 methylation. a) A region of CpGs within exon 2 of IRX1 were significantly hypermethylated in the cocaine group as per RRBS analysis (n= 25 per group), b) however this effect did not replicate in a second, independent cohort (n= 20 controls; n=15 cases). Data represented as mean ± s.e.m. *** q< 0.05



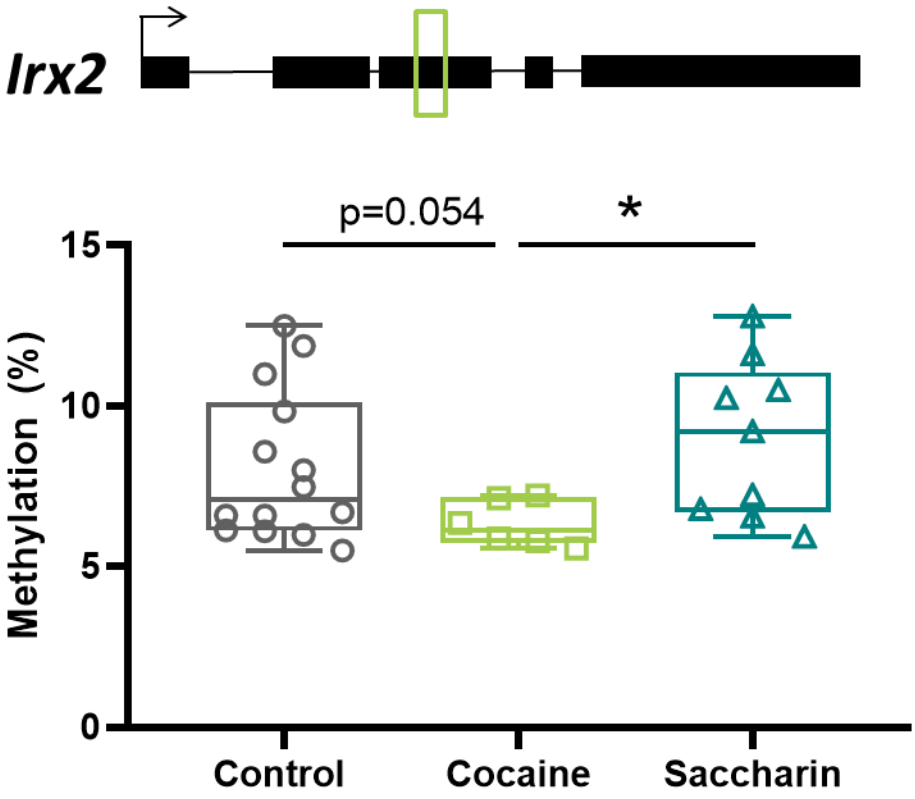
Supplemental Figure 5. Fluorescence Activated Nuclei Sorting (FANS) in post-mortem caudate nucleus. a) An example sorting plot, with gates indicated separating neuronal (NeuN+) from non-neuronal (NeuN-) nuclei. b) Post-sort purity analysis supports the accuracy of sorting based on NeuN. c) The ratio between nuclei types, d) number of NeuN+ nuclei, e) and number of NeuN- nuclei were not different between cases (n=22) and controls (n=21).



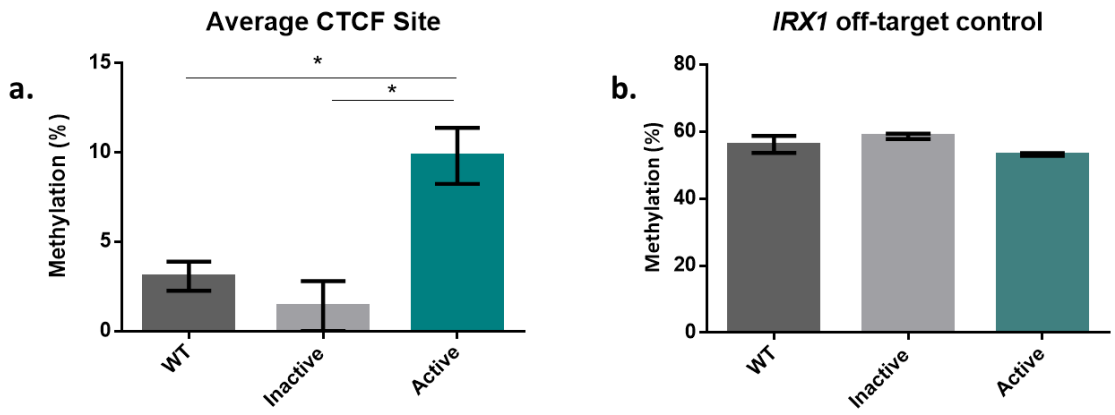
Supplemental Figure 6. No change in *IRX2* methylation in non-neuronal nuclei. a) average methylation across exon 3 did not differ between non-neuronal (NeuN-) nuclei from the cocaine and control groups (n=24 controls and 23 cases; discovery cohort), b) nor did methylation across the 4 CpGs within the putative CTCF binding site. Bar data represented as mean



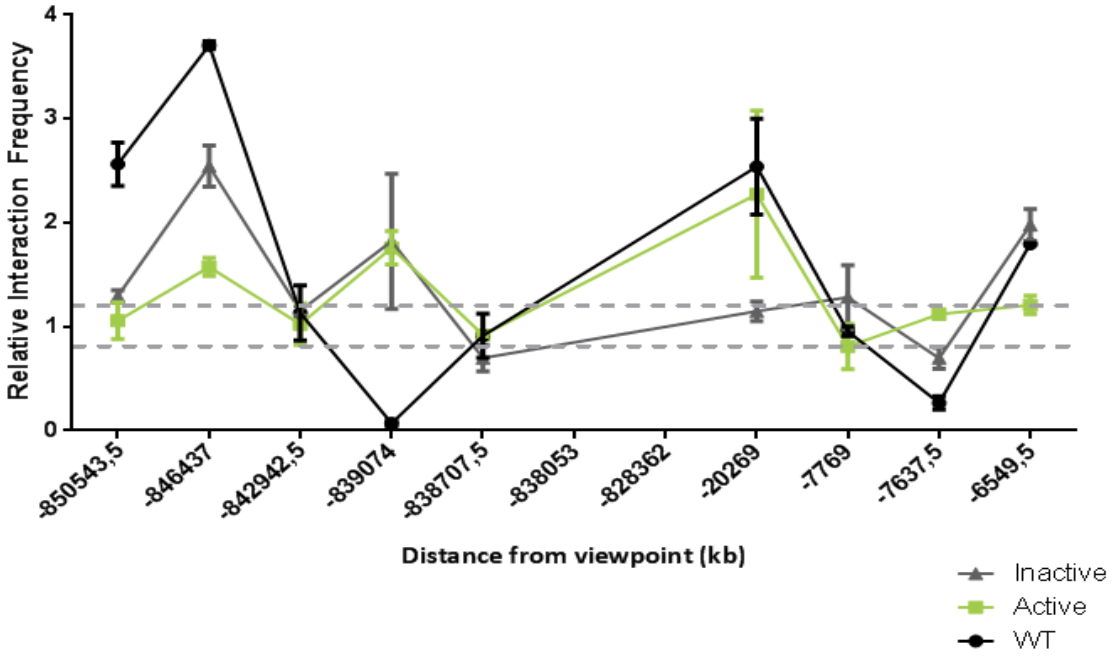
Supplemental Figure 7. Acquisition of reward-consuming behavior in mice. Animals trained to self-administer a) cocaine (n=8) and b) saccharin (n=11) consistently performed active nosepokes to receive reward injections over the course of 10 days, compared to responses from saline-injected animals (n=16).



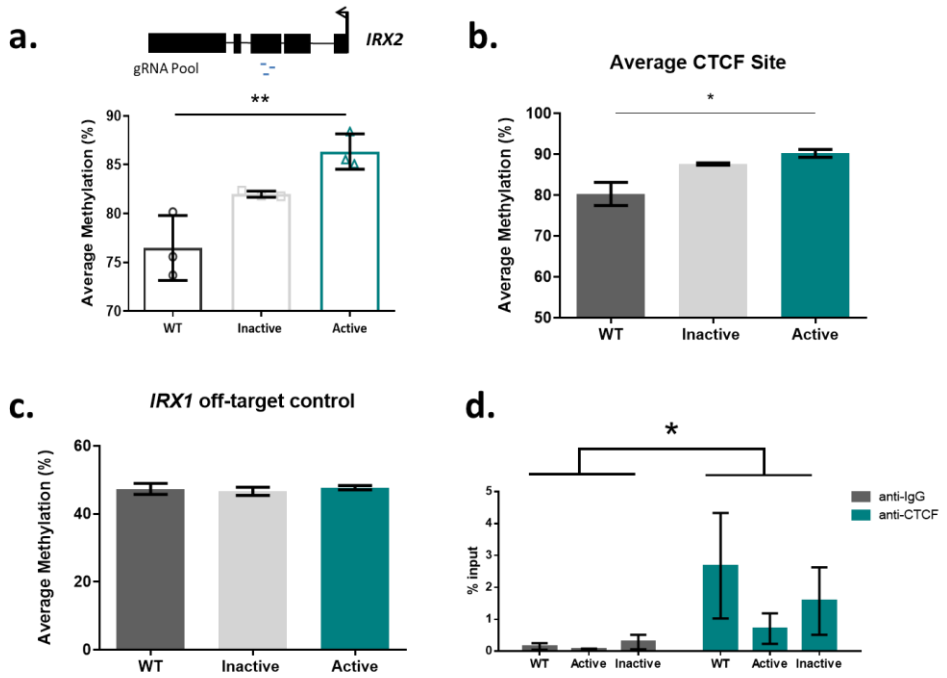
Supplemental Figure 8. Exon 3 of Irx2 is hypomethylated in cocaine self-administering mice. Cocaine self-administering mice (n=6) had significantly less methylation, on average, within a homologous CTCF binding site in *Irx2,* in caudate putamen homogenate samples, when compared to control (n=14) and saccharin (n=9) groups. * $p < 0.05$



Supplemental Figure 9. dCas9 targeted methylation was specific to IRX2. a) Average methylation of putative CTCF site RENcells (n=3) is increased compared to inactive and control cells, while b) off-target methylation of IRX1 was absent. Data represent mean ± s.e.m. *p < 0.05



Supplemental Figure 10. Altered chromatin landscape of dCas9-DNMT3A edited HEK293 cells. Data represent mean ± s.e.m.



Supplemental Figure 11. Confirmation of CTCF binding to putative binding site in *IRX2* exon 3. a) Transfection of an active dCas9-DNMT3A construct, along with a pool of 3 guide RNA constructs significantly increased methylation of *IRX2* exon 3 in HEK293 cells, compared to transfection with an inactive construct or wildtype controls (n=3 replicates of 5 x 10^6 cells per group). b) methylation of the putative CTCF site HEK293 is increased compared to control cells, while c) off-target methylation of IRX1 was absent. d) CTCF binds more frequently to exon 3 than a non-specific IgG antibody control, across all experimental groups**.** Data represent mean ± s.e.m. *p < 0.05, **p <0.01

CHAPTER III: DNA METHYLATION OF A NOVEL REGULATORY ELEMENT WITHIN THE TYROSINE HYDROXYLASE GENE (*TH*) IS DYSREGULATED BY CHRONIC COCAINE DEPENDENCE IN THE HUMAN STRIATUM

DNA methylation of a novel regulatory element within the tyrosine hydroxylase gene (*TH*) is dysregulated by chronic cocaine dependence in the human striatum

Kathryn Vaillancourt[1,2], Gang G Chen[1], Laura Fiori[1], Gilles Maussion[3], Volodymyr Yerko[1], Jean-François Théroux[1], Carl Ernst[7], Benoit Labonté[4], Erin Calipari[5], Eric J. Nestler[6], Corina Nagy[1,7], Naguib Mechawar[1,7], Deborah C. Mash[8], Gustavo Turecki[1,2,7]

[1]McGill Group for Suicide Studies, Douglas Hospital Research Center, Verdun, QC, Canada; [2]Integrated Program in Neuroscience, McGill University, Montreal, QC, Canada; [3]Department of Neurology and Neurosurgery, Montreal Neurological Institute, Montreal, Quebec, Canada; [4]Centre de Recherche Cervo, Université Laval, Québec, QC, Canada; [5]Departments of Pharmacology, Molecular Physiology and Biophysics, Psychiatry and Behavioral Sciences; Vanderbilt Center for Addiction Research; Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA; [6]Nash Family Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [7]Department of Psychiatry, McGill University, Montreal, QC, Canada; [8]Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, USA

Currently under review at *eLife.* Modified for this thesis.

Abstract: Cocaine dependence is a chronic, relapsing disorder caused by lasting changes within the brain's reward circuitry. Methylome-wide studies have identified multiple biological pathways that contain altered DNA methylation dynamics in the striatum of rodents; however, there is a lack of information about the involvement of DNA methylation in cocaine dependence in humans. Here, we generated genome-wide methylomic profiles for the nucleus accumbens, part of the ventral striatum, using human postmortem brains from a cohort of cocaine dependent and healthy controls (n=25 per group). In particular, we have identified hypermethylation of a cluster of CpGs within the gene body of tyrosine hydroxylase gene *(TH),* which contains a putative binding site for the early growth response 1 (ERG1) transcription factor. We replicated this finding in striatal tissue from an independent cohort (n=18 per group) and found it to be specific to striatal neuronal nuclei. Furthermore, this locus demonstrates enhancer activity in vitro, which is attenuated by methylation and enhanced by EGR1 overexpression. This is the first time that epigenetic dysregulation has been associated with cocaine dependence with respect to dopaminergic signaling and represents an important advancement in our understanding of drug neurobiology in humans.

Keywords: cocaine, addiction, epigenetics, DNA methylation, tyrosine hydroxylase, striatum, nucleus accumbens, caudate nucleus, psychiatry, post-mortem brain

1. Preface

Our previous study demonstrated that cocaine dependence is associated with widespread alterations in DNA methylation in the human caudate nucleus. We place particular focus on a novel intragenic regulatory region within *IRX2* that was less methylated in neurons of the cocaine group, contained a CTCF binding site, and altered gene expression through its control of three-dimensional chromatin looping. The goal of this study was to follow-up on this work with a similar, methylome-wide investigation in the nucleus accumbens, a striatal subregion that is heavily involved in the motivational aspects of drug seeking and has serial connectivity with the caudate in addicted subjects. One of the most significantly differentially methylated loci from this analysis was an intragenic region in tyrosine hydroxylase (*TH*), a gene with important implications for cocaine neurobiology. In fact, we found the same cluster of CpGs to be hypermethylated in our caudate dataset and explored the specificity of these findings using an independent cohort of tissue, as well as FANS sorted nuclei. We found this region to be affected in either the accumbens or caudate in each comparison and although we saw no change in *TH* gene expression, we used an in vitro enhancer model to show that the region has enhancer potential. Moreover, the regulatory potential of this locus is sensitive to cocaine-related changes in methylation as well as EGR1 binding, a protein whose gene we found to be overexpressed in both brain regions fin the cocaine group. These results suggest that there is remarkable similarity in the characteristics of cocaine-related methylation in both brain areas, although the majority of loci are independent. Furthermore, this study implicates a rare, newly discovered subtype of striatal GABAergic interneurons in drug use disorders.

2. Introduction

Chronic cocaine exposure — and the behavioral symptoms that accompany the cocaine dependence phenotype — are associated with long lasting changes in brain biology. Long-term dysregulation of signalling in downstream targets of midbrain dopamine projections accompanies compulsive drug seeking behavior and withdrawal (Nestler and Lüscher 2019; Tomasi et al. 2010; Volkow et al. 2006). Accordingly, regions of the mesocorticolimbic neurocircuit, including the dorsal (caudate nucleus) and ventral (nucleus accumbens) striatum, have been actively studied in humans and animal models. In humans, serial connectivity between the caudate and accumbens increases as patients transition to an addicted state, and widespread transcriptomic changes are associated with chronic exposure paradigms in rodents (Albertson et al. 2004; Belin and Everitt 2008; Walker et al. 2018). Epigenetic mechanisms, including histone post translational modifications and cytosine methylation, have been implicated as mediators of the impact of these relationships, and DNA methylation in particular has been shown to be necessary for the development of drug seeking behaviors (LaPlant et al. 2010).

Recently, we have shown that chronic cocaine dependence in humans is associated with genome-wide alterations in DNA methylation within the caudate nucleus (Vaillancourt et al. 2020). In the present study, we complement these findings with methylome-wide profiles of cocaine dependence in human post-mortem nucleus accumbens and provide evidence for novel epigenetic regulation of the tyrosine hydroxylase gene (*TH*). We have found a region within *TH,* which contains portions of exons 8 and 9 as well as the intron between them, and which is hypermethylated in both striatal subregions and is modulated by changes in methylation and transcription factor availability in vitro.

3. Results

DNA methylation profile of cocaine dependence in human nucleus accumbens

We investigated nucleus accumbens tissue (Figure 1a) dissected from 25 individuals with chronic cocaine dependence and 25 psychiatrically healthy, drug-naïve controls (Supplementary Table 1) and performed reduced representation bisulfite sequencing (RRBS) to identify methylome-wide alterations associated with case status (Chen et al. 2014). This strategy resulted in similar sequencing depth, bisulfite conversion rates, and CpG coverage between groups (Supplementary Table 2). Our analysis identified 4814 clusters of CpGs that passed filtering, including 145 differentially methylated regions (DMRs) with a Benjamini-Hochberg corrected q-value < 0.05 when comparing cases and controls (Supplementary Table 3), and the majority of DMRs were located within 5 kb of an annotated gene (113; 77.9%). Although DMRs were distributed across all autosomes (Figure 1b), and varied in cluster size and strength of differential methylation, there were significantly more hypermethylated DMRs than hypomethylated (Chi squared Goodness of Fit = 4.235; $p < 0.05$; Figure 1c, Table 1), which is analogous to our previous findings in the dorsolateral caudate nucleus (Vaillancourt et al. 2020) and is consistent with animal studies of cocaine self-administration (Baker-Andresen et al. 2015). Next, we overlapped our list of DMRs with datasets containing known genomic and epigenomic features and found that all identified clusters, including those which were differentially methylated, overlapped primarily with CpG islands (data not shown) and intragenic regions, including introns, exons, and intron-exon boundaries (Supplementary Figure 2a; Figure 1d). Interestingly, when we compared our dataset to ChromHMM predicted epigenome states, we found a significant enrichment of regions overlapping with active transcription start sites in the DMRs compared to the list of all clusters (FDR < 0.05; Figure 1e, Supplementary Figure 2b). Finally, we searched for potential functional commonality between the genes nearest to the DMRs using PANTHER gene ontology analysis (Thomas et al. 2003). Although it is understood that the nearest gene to a regulatory site is not always its primary target (Schoenfelder and

Fraser 2019), we still found the genes nearest to the DMRs to be enriched for DNA binding proteins and transcription factors (FE=5.74; FDR < 0.05).

Since chronic cocaine seeking in animals is associated with changes in the expression of DNA methyltransferase genes (LaPlant et al. 2010; Tian et al. 2012; Wright et al. 2015), we first investigated the expression of these genes. We found that *DNMT3A* was significantly increased in the nucleus accumbens of the cocaine group compared to controls (t=2.96, df=41, p=0.005; Supplementary Figure 1a), as was the case in the caudate (Vaillancourt et al. 2020). The expression of the other two methyltransferase genes, *DNMT3B* and *DNMT1*, was not associated with case status (ps > 0.05; Supplementary Figure 1b and 1c).

Notably, a cluster of CpGs within the body of the tyrosine hydroxylase gene (*TH*) was among the topmost hypermethylated loci in our analysis (9.17%, Table 1). Since TH is a necessary component of dopamine synthesis and has a known relationship with reward-driven behaviors (Kaminer et al. 2019; Logan et al. 2019) we were immediately struck by its potential importance. Interestingly, *TH* was also identified among the short list of DMRs that were shared between the current study and our previous work in the caudate nucleus (orange data points Figure 1b, Table 2), where the exact same cluster of CpGs was differentially methylated in the same direction, and similar to a magnitude (9.09%). The *TH* DMR is comprised of 3 CpGs within a CpG island in the eighth and nineth exon of *TH,* and we chose to further investigate this locus in the context of human cocaine neurobiology.

Tyrosine hydroxylase hypermethylation is related to cocaine dependence in the caudate and accumbens

In order to validate and extend our findings, we performed deep bisulfite amplicon sequencing (Chen et al. 2017) on separate aliquots of genomic DNA from a larger sample set including the discovery

cohort, and again found increased methylation to be associated with cocaine dependence in the caudate (Mann-Whitney test, $p < 0.05$, Supplemental Figure 3a), with no significant difference between groups in the nucleus accumbens ($p > 0.05$, Supplemental Figure 3b). Furthermore, since methylation of individual CpGs within an island tends to be related (Barrera and Peinado 2012), we decided to extend the amplicon beyond the original 3 CpGs and interrogated the methylation status of an extended group of 27 cytosines within the intragenic CpG island (Figure 2a). Here, we found 5.8% more methylation across all CpGs in the caudate nucleus in the cocaine group (Mann Whitney test, $p < 0.01$; Figure 2b), and a similar trend in the nucleus accumbens with a 3.2% increase in methylation (Mann Whitney test, $p = 0.058$; Figure 2c). We generated RRBS data from the striatum of mice who were trained to self-administer cocaine and used it to compare the methylation status of a cluster of 11 CpGs within the eighth exon of *Th* (Supplemental Figure 4a). We found that chronic cocaine exposure and seeking was related to 4.8% more methylation at *Th* within the nucleus accumbens ($t=4.115$, $df=16$, $p < 0.001$, Supplemental Figure 4b), and no significant increase in the caudate-putamen ($p > 0.05$, Supplemental Figure 4c).

_TH_ hypermethylated region includes an EGR1 binding motif, and is specific to neuronal nuclei

Next, to investigate the potential functional implications of increased methylation at our DMR, we used published chromatin immunoprecipitation data (ENCODE, (Bernstein et al. 2012)) to search for DNA binding proteins. One such factor, early growth response protein 1 (EGR1), was found to bind within the *TH* gene in multiple cell models and has a putative consensus motif (5'-CGGCCCCCCGGC-3', Supplemental Figure 5; Figure 3a, (Zandarashvili et al. 2015)) within the cocaine-associated DMR. This potential binding site contains CpGs at both ends (at hg 19_ch11:2188079 and 2188087), which demonstrate the expected hypermethylation in the caudate nucleus (Two-way repeated measures ANOVA, main effect of Group, $F(1,49)=8.914$, Sidak's multiple comparisons $p < 0.01$, Figure 3b), but not in the nucleus accumbens (Two-way repeated measures ANOVA, $F(1,49)=0.65$, $p > 0.05$, Figure 3c).

We then sought to strengthen our findings through replication and investigated tissue from an independent cohort of psychiatrically healthy controls and individuals who died with chronic cocaine dependence (n=18 per group, Supplemental Table 4). Using an identical bisulfite amplicon sequencing strategy, we found no differences in the caudate nucleus and significantly higher methylation across the putative EGR1 site in the nucleus accumbens in this cohort (Two way repeated ANOVAs, Fs(1,33)=2.292, 5.745 ps = 0.1399, 0.0224, respectively, Figure 3d, e). Although the variability within each dataset reduced our power to detect significant differences, there is a trend towards increased methylation of both CpGs within the EGR1 binding site between two brain regions and across both cohorts.

DNA methylation patterns are cell-type specific and are related to distinct cellular functions (Kozlenkov et al. 2015; Lister et al. 2013; Luo et al. 2017), so we used fluorescence-activated nuclei sorting (FANS) to separate neuronal from non-neuronal nuclei in new dissections of caudate and accumbens tissue from the discovery cohort. We separated intact nuclei from cellular debris using the DRAQ5 DNA stain, and neuronal from non-neuronal nuclei based on fluorescence with the NeuN neuronal marker (Figure 3f). We hypothesized that DNA methylation within the flanking CpGs of the EGR1 binding site would be significantly higher in neuronal nuclei from the cocaine group than in controls, and that this effect would be absent in non-neuronal nuclei. Indeed, in the caudate nucleus, we observed increased methylation in neuronal, but not non-neuronal, nuclei from the cocaine group (Two-way repeated measures ANOVAs, Fs (1,40) = 0.75, 0.08 and ps > 0.05 Figure 3g, 3h). We found no effect of group on DNA methylation in either nuclear fraction in the nucleus accumbens (Two-way repeated measures ANOVAs, Fs (1,40) = 7.410, 0.1581 and ps=0.0210, 0.7621; Supplemental Figure 6).

Exon 8/9 of *TH* has enhancer activity that is regulated by DNA methylation and EGR1 expression

Since the regulatory potential of this genomic region has yet to be reported, we asked whether the segment containing part of exon 8 and 9, along with the intron between them, can act as an

enhancer for *TH* gene expression. We performed an enhancer luciferase assay to assess the regulatory potential of a 398 bp fragment of our differentially methylated region with *TH.* We cloned the fragment, into a pCpG-free vector upstream of the *EF1α* promoter and transfected methylated or unmethylated constructs into HEK293 cells (Figure 4A). After 24 hours, we quantified luciferase activity and found a highly significant effect of group (One-way ANOVA, $F=169.95$; $df=4$; $p=1.37\ 10^{-14}$, Figure 4B). The inclusion of the *TH* fragment resulted in a 7.5-fold increase in luciferase activity compared to the empty vector (Empty vector vs. *TH*, Tukey's multiple comparisons; $p < 0.0001$, Figure 4B). Furthermore, when the fragment was fully methylated prior to transfection, the enhancer activity of the region was effectively lost, where luciferase expression was not significantly different from the empty vector group (Empty vector vs. Methylated *TH*, Tukey's multiple comparisons; $p > 0.05$, Figure 4B). This strongly suggests that the cocaine-related DMR in the *TH* gene has regulatory potential, which is attenuated by DNA methylation.

We further hypothesized that the enhancer activity of this locus would be increased by the presence of EGR1, so we co-transfected cells with one of the *TH* enhancer vectors described above, and either an EGR1 overexpression vector or a negative control. We found that excess EGR1 greatly increased enhancer activity (*TH* vs *TH* +EGR1, Tukey's multiple comparisons; $p < 0.0001$, Figure 4B). Importantly, when we co-transfected the methylated *TH* plasmid with the EGR1 overexpression vector, we found a significant decrease in luciferase activity compared to the unmethylated TH plasmid, either with or without EGR1 overexpression (Methylated TH + EGR1 vs TH + EGR1 or vs TH, Tukey's multiple comparisons; $p < 0.0001$, Figure 4B). Together, these results suggest that the intragenic region within *TH* may serve as a regulatory element to nearby promoters, and that it is responsive to upregulation by EGR1 and downregulation by DNA methylation.

Finally, given the apparent relationship between DNA methylation and EGR1 on the regulatory potential of this genomic region, we asked whether *TH* or *EGR1* mRNA expression was dysregulated with

cocaine dependence in the human striatum. We found significantly more *EGR1* expression in both the caudate nucleus (Mann Whitney U = 144, p = 0.008, Figure 5a) and the nucleus accumbens (t=2.38, df=32.9, p=0.023, Figure 5b) in individuals who had chronic cocaine dependence compared to controls, which is consistent to what has been found in animal models of psychostimulant exposure (Moratalla, Robertson, and Graybiel 1992). Conversely, we found no significant differences in *TH* expression in either brain region (ps > 0.1, Figure 5c, d). Potential explanations for these data are offered in the next section.

4. Discussion

This study has generated the first cocaine dependence methylome in human nucleus accumbens and fills a vital gap in our knowledge of how chronic cocaine exposure and addiction impact brain cell biology. We have identified over 100 DMRs that are associated with cocaine dependence in humans, including an enrichment for those overlapping with intragenic regions and active transcription start sites. This mirrors our recent findings in the human caudate nucleus (Vaillancourt et al. 2020) and fits well with the knowledge that intragenic regions, including enhancers, are important to brain health. For example, methylation at enhancers plays an important role in neuronal cell identity, and genetic polymorphisms in enhancer regions confer increased risk for multiple psychiatric disorders (Hannon et al. 2019; Kozlenkov et al. 2015). Additionally, gene body methylation has been shown to regulate gene expression through the regulation of alternative promoters and mRNA splice variants, which have both been implicated in psychiatric disorders including cocaine dependence (Feng et al. 2014; Gandal et al. 2018; Maunakea et al. 2010, 2013). It stands to reason that epigenetic modifications, including DNA methylation, within these regions might play an important role in disease-related brain functioning as well.

In addition, we found an over-representation of hypermethylated loci within the striatum, which adds to decades of work in animal models using multiple drug delivery paradigms and time-courses (Vaillancourt et al. 2017). For example, cocaine self-administration in rodents results in more hypermethylation than hypomethylation in some downstream targets of midbrain dopamine projections (Baker-Andresen et al. 2015; Fonteneau et al. 2017). In addition, expression of the *DNMT3a* de novo methyltransferase is induced by cocaine exposure and can be maintained after extended periods of withdrawal (LaPlant et al. 2010). Furthermore, these phenomena combined, in the nucleus accumbens, appear to be necessary for the maintenance of compulsive drug seeking behaviors in rats (Massart et al. 2015). In human post-mortem tissue, we found *DNMT3A* to be increased after chronic cocaine dependence in the caudate nucleus of humans (Vaillancourt et al. 2020), and the present study recapitulates this finding in the nucleus accumbens. These results, along with the distribution of DMRs, suggests that there are genome-wide alterations in experience-dependent DNA methylation associated with chronic cocaine dependence.

When comparing DMRs between brain regions, we found the exact same cluster of CpGs within exon 8 of *TH* to be hypermethylated in the caudate nucleus and nucleus accumbens of human patients; our expanded analysis included 27 CpGs that spanned a portion of both exons 8 and 9, as well as the intronic region that separates them. We found no significant difference in *TH* mRNA expression in either the caudate nucleus or the nucleus accumbens in humans. There are several possible explanations for these negative data, in particular, the lack of cellular resolution. *TH* gene expression dominates in midbrain dopamine neurons that innervate striatal regions, while relatively low levels of *TH* mRNA are found in these downstream targets. It is assumed that most of this mRNA is present in dopaminergic nerve terminals that innervate striatum (Gervasi et al. 2016), although a single nucleus transcriptome study in rats and an electrophysiology study in mice identified sparse cell clusters expressing *Th* in the nucleus accumbens; such transcript levels were unchanged following cocaine exposure (Ibáñez-Sandoval

et al. 2010; Savell et al. 2020). It is important to note, however, that this study examined changes after a single exposure to cocaine and therefore does not directly compare to the prolonged behavioral changes associated with a dependent phenotype. Additionally, since both *EGR1* expression and DNA methylation within *TH* are increased in human subjects, *TH* transcript levels may be kept in check through a combination of opposing mechanisms. In any event, future studies using single-cell technologies and dependence-like behavioral models are likely required to identify changes in mRNA expression, protein levels or protein activity in the striatal regions studied here. Most importantly, however, our discovery of a novel enhancer role for the designated region of *TH* raises the possibility that its methylation and EGR1 binding controls the expression of other genes, spatially connected to this locus, in striatal regions.

TH codes for tyrosine hydroxylase, the rate-limiting step in the biosynthesis of dopamine within the terminals of presynaptic cells. It has been well documented that as patients transition from recreational cocaine use to physiological dependence and addiction, they undergo lasting changes in dopamine signalling within the mesocorticolimbic circuitry (Letchworth et al. 2001; Volkow et al. 1999). Furthermore, although transcription and protein immunoreactivity of TH is known to increase after extended cocaine exposure in the rodent midbrain (Beitner-Johnson, Guitart, and Nestler 1991; Logan et al. 2019; Masserano et al. 1996), attempts to measure the effects of cocaine on TH in downstream targets have yielded mixed results. TH immunoreactivity is upregulated in the central amygdala after 45 days of withdrawal from cocaine self administration in rats, however, there appears to be no change in protein levels in the prefrontal or orbitofrontal cortex after cocaine conditioned place preference (Grimm, Shaham, and Hope 2002; Hámor et al. 2020). In the dorsal striatum, chronic cocaine injection increases TH protein levels, while the number of TH positive varicosities in the nucleus accumbens has been reported to be both increased and decreased after cocaine exposure, depending on tissue sub-sectioning and length of withdrawal (Balda, Anderson, and Itzhak 2009; Schmidt et al. 2001; Todtenkopf,

De Leon, and Stellar 2000). Interestingly, the decreased immunoreactivity that is found in the nucleus accumbens shell in rats after 12 days of self-administration can be reversed by extinction training and appears to be specific to cocaine and not other appetitive reward stimuli (Schmidt et al. 2001). This suggests that levels of TH protein abundance and activity in the nucleus accumbens have an important role in cocaine-seeking behavior, however, it is still unclear whether the protein in these studies is located within afferent dopaminergic projections or within striatal cell bodies themselves.

In this study, we found a variable increase in methylation within the gene body of *TH* in the human striatum, that may be specific to neuronal nuclei in the caudate nucleus. The variability within our data, along with our lack of sensitivity to detect transcriptional dysregulation is highly suggestive of these findings occurring in a small proportion of cells. Indeed, in a recent single-nucleus transcriptomic survey of the rat nucleus accumbens, *Th* transcripts were detected in less than 30% of nuclei, and mouse data show very few instances of *Th* detection in single-cell transcriptome experiments (Saunders et al. 2018; Savell et al. 2020). Canonically, the mammalian striatum is composed primarily of D1- and D2-expressing medium spiny projection neurons (MSNs, 75-95%), as well as at least 4 subtypes of inhibitory interneurons (Lobo et al. 2006). Among them, are a rare group of TH-expressing interneurons (THINs, approximately 0.04%) that have distinct electrophysiological properties (Ibáñez-Sandoval et al. 2010). THINs do not appear to release dopamine, and do not co-localize with other proteins involved in monoamine synthesis (Xenias et al. 2015). Instead, these cells exhibit strong, GABAergic inhibition of MSNs in response to thalamic input, and may have a distinct role in striatum-related behaviors; selective ablation of THINs results in deficits in specific goal-directed behaviors (Kaminer et al. 2019; Xenias et al. 2015). Although the present study represents the first step towards identifying cell-type specific methylation patterns associated with chronic cocaine dependence in humans, the findings with respect to *TH* should be further examined in specifically relevant datasets including those targeting THINs.

Additionally, although we are the first to report increased *EGR1* mRNA expression following cocaine dependence in humans, this transcription factor is known to have an important role in the neurobiology of cocaine in rodents. It is an immediate early gene which is induced in both the dorsal and ventral striatum via dopaminergic signalling upon initial exposure to cocaine, but it is also necessary for context-related drug cues (Bhat, Cole, and Baraban 1992; Fritz et al. 2011; Moratalla et al. 1992). EGR1 is necessary for the development of cocaine conditioned place preference, and its expression is re-induced upon exposure to drug-related environments, even after prolonged abstinence (Hearing, See, and McGinty 2008; Valjent et al. 2006). Our data supports the notion that EGR1 is involved in cocaine-induced neuroadaptation, and we extend its role to include activation of a novel regulatory element within *TH*, thereby connecting two previously distinct pathways of dysregulation.

The use of post-mortem human tissues is invaluable in understanding fundamentally human disorders, and although we support our hypotheses with in vitro studies, this work is not without limitations. Given the difficulty in tissue dissection, we are unable to disentangle the effects the core versus the shell of the nucleus accumbens in our data, although each subregion is known to have distinct roles in drug seeking behavior in animals (Dumitriu et al. 2012). Furthermore, we have shown that the hypermethylation of the EGR1 binding site is specific to neuronal nuclei in the caudate nucleus, however given the magnitude and variability of this effect, we suspect this signal to be coming from a small proportion of cells. As single-nucleus DNA methylation technologies advance so will our understanding of the importance of specific, TH-expressing cell types in cocaine neurobiology.

Our study is the first to provide evidence of DNA methylation dysregulation in the TH gene that is associated with cocaine dependence; however, given the nature of post-mortem research, we are unable to definitively conclude whether this dysregulation is a direct consequence of the addicted phenotype. It is possible that altered *TH* methylation prior to drug exposure could predispose some individuals to develop cocaine dependency, but we are unable to test this hypothesis in human subjects.

Future studies should make use of in vitro epigenome editing, using targeted DNA methylation in the striatum of living animals, to investigate the behavioral outcomes of *Th* hypermethylation.

This study has identified a novel regulatory element within *TH* that contains a putative EGR1 binding site and is differentially methylated in the striatum of individuals with cocaine dependence. Although we are the first to demonstrate in vitro enhancer activity of this locus, a downstream section of the gene beginning with exon 9 (part of which, is included within our region of interest) has been identified as a bivalent promoter, and displays a local enrichment of the enhancer-associated histone marks H3K27ac and H3K4me1 in the Roadmap Epigenome Consortium dataset (http://epigenomegateway.wustl.edu, (Kundaje et al. 2015)). Interestingly, these features are annotated within the samples from the human striatum, but not the substantia nigra, which is a midbrain source of dopamine production. Although more experiments are needed, this phenomenon in addition to the findings presented here, suggests a specific role of this locus in cocaine-related neurobiology in the nucleus accumbens and caudate nucleus.

5. Methods

Subjects

All the methods in this study were approved by the Douglas Hospital Research Ethics Board, and human post-mortem tissues were acquired from the Brain Endowment Bank at the University of Miami Miller School of Medicine, where autopsy and tissue handling was performed in accordance to the established standards. For the discovery cohort, nucleus accumbens and dorsolateral caudate nucleus tissue was dissected from 25 subjects who had long term histories of cocaine dependence (as determined by licenced clinicians), and who died from cocaine-related complications as determined by forensic pathology and tissue toxicology. These subjects ("cases") were selected based on the absence of illicit drug toxicology other than cocaine, and the absence of other psychiatric diagnoses as

determined by medical records and reports from next of kin. Additionally, 25 drug-naïve, psychiatrically healthy subjects who died from accidental or natural causes, were selected as "controls". The replication cohort was selected from independent "cases" and "controls" using the same selection criteria (n=18 per group), and nucleus accumbens and caudate nucleus tissue was dissected as described above. All subjects in this cohort were male, which is reflective of the opportunistic composition of the brain samples available at autopsy, and no *a priori* power analyses were performed in relation to sample sizes, due to the rarity of these samples.

Reduced representation bisulfite sequencing (RRBS)

Library Preparation

We used 20 mg of frozen nucleus accumbens tissue from each subject in the discovery cohort to extract genomic DNA (gDNA), using Qiagen DNA MiniKits as per manufacturer's instructions. Next, we digested 1ug of gDNA overnight with the *MspI* restriction enzyme and completed RRBS library preparation as described elsewhere (Chen et al. 2014). We converted purified libraries with the EpiTect fast bisulfite conversion kit (QIAGEN, Cat# 59824) according to the standard protocol, and indexed them through PCR amplification before sequencing.

Sequencing and Bioinformatic Processing

Indexed libraries were sequenced on the Illumina HiSeq 2000 platform at the Genome Quebec Innovation Center (Montreal, Canada) with 50bp single end sequencing. All bioinformatic processing and quality control was performed in-house, as described (Chen et al. 2014). We calculated the mean bisulfite conversion efficiency using the ratio of reads containing T to C at the unmethylated cytosine position that was added during the end-repair step of library construction for each library, and then averaged across group (case and control).

Differential methylation analysis

We filtered out CpGs that were not present in at least 25 subjects from both cases and controls and had less than 5X coverage across libraries. Next, we grouped CpGs within 50 bp of each other into clusters using the bumphunter 3.5 package for R, and removed those that had a standard deviation of less than 5% methylation across all subjects (i.e., irrespective of status) to reduce unnecessary comparisons between stably methylated regions. Each cluster contained at least 2 CpGs, without an upper limit to cluster size, resulting in 32 535 clusters moving forward. We analysed each cluster using a general linear model with status (case or control) as a fixed factor and covariates including ethnicity, age, smoking status, and ethanol toxicology. We treated CpGs independently in each cluster but investigated only those clusters that had a Bejamaini-Hochberg FDR corrected p-value <0.05 and which were <0.05 when calculating a single mean from all CpGs per individual.

Enrichment Analyses

We used the *annotatr 1.10.0* package in R to annotate all CpG clusters and DMRs to their genomic context, CpG island proximity, and predicted ChromHMM chromatin state using (Cavalcante and Sartor 2017). We used the 15-core marks from ChromHMM that were generated using human striatum tissue and calculated enrichment q values for DMRs against all CpG clusters using the *LOLA* algorithm(Sheffield and Bock 2016) (cite ChromHMM).

For gene ontology, we matched each DMR to its nearest Refseq gene and used over representation tests in the gene list analysis functions of the PANTHER classification system (www.pantherdb.org). We compared the full DMR list to all human genes with respect to molecular function, biological processes and cellular components and *p*-values were calculated using Fisher's Exact tests with FDR correction.

RNA sequencing

We used RNeasy Lipid Tissue Kits (Qiagen) to extract RNA from 100mg of nucleus accumbens tissues, according to the manufacturer's standards. The RNA integrity number (RIN) for controls was, on average, 8.1 ± 0.6 and 8.1 ± 1.2 for cases, as measured with RNA bioanalyzer chips (Agilent). Aliquots of 100ng/ul of RNA were sent for library preparation and sequencing at the Broad Institute (Cambridge, MA) and libraries were prepared using a standard non-strand specific protocol (Illumina TruSeq), including poly-A selection, and multiplexed for 50bp paired end sequencing on the Illumina HiSeq 2000 platform, and alignment was performed and reads per kilobase per million (RPKM) were calculated as previously described(GTEx Consortium 2015).

Data was extracted for the three methyltransferase genes (DNMT1, DNMT3A, and DNMT3B), and RPKM values were averaged across groups, after removing statistical outliers (ROUT Q=1%). Cases and controls were compared using two-tailed student's t-tests.

Cocaine self-administration in mice

Self-Administration

Male C57/B6 mice (n= 10 per group) were anesthetized with ketamine (100 mg/kg) and xylazine (10 mg/kg) implanted with chronic indwelling jugular catheters and trained for i.v. self-administration as previously described(Thomsen and Caine 2007). The catheter tubing was passed subcutaneously from the back to the jugular vein and 1.2 cm of tubing was inserted into the vein and secured with silk suture. Following surgery, animals were singly housed, and allowed to recover for 48-72 hours. Each animal was maintained on a reversed light cycle (7:00am lights off; 7:00 pm lights on) and all self-administration procedures occurred during the active/dark cycle. Sessions were two hours in length and animals self-administered cocaine (0.5 mg/kg/inj over 3 sec) on a fixed-ratio 1 schedule of administration. At the beginning of each session a house light was illuminated signifying the availability of drug. After recovery from surgery, the animals were placed in a drug self-administration operant chamber where

they were connected to a drug-intake line. The operant boxes are equipped with two retractable levers; depression of the “active” lever resulted in a drug infusion, whereas no infusion occurred by pressing the “inactive” lever. Concurrent with the start of each injection, the lever retracted, the house light was turned off, and a stimulus light was activated for 5 seconds to signal a time-out period. Under these conditions, animals acquired a stable pattern of intake within 1 to 5 days. For self-administering animals, acquisition (Day 1) was counted when the animal reached 70% responding on the active lever and 10 or more responses. Following acquisition, the animals were given access to a cocaine-paired lever for 120 minutes per day for a period of 10 days. Control animals underwent the same experimental procedures but had access to a saline-paired lever.

Reduced Representation Bisulfite Sequencing in mice

After training, animals were sacrificed, and brain tissue was removed, and flash frozen at -80°C until dissection. Caudate-putamen and nucleus accumbens tissue was dissected from frozen tissue, and DNA was extracted using Qiagen DNA Minikits as per manufacturer’s protocol. RRBS libraries were prepared as above, using 2ng of genomic DNA from each animal. Each library was individually barcoded, and pooled for 50 bp SE sequencing, 4-5 libraries per lane, on the Illumina HiSeq 2000 platform at the Genome Quebec Innovation Center. Bioinformatic processing was done as above, except reads were aligned to the mouse genome (GRCm38/mm10). Rather than methylome-wide analysis, we searched the computed clusters for any that aligned with the Th gene (chr7:142,892,752-142,900,014) and found a single cluster within exon 8 that contained 11 CpGs, and was detected in at least 80% of sequenced libraries. After removing statistical outliers using ROUT tests (Q=1%), we performed two tailed Student’s t-tests between control and cocaine animals, separately for the dorsal and ventral striatum.

Targeted bisulfite amplicon sequencing

Library Preparation

We used Methyl Primer Express Software v1.0 (Applied Biosystems, CA, USA) to design redundant pairs of bisulfite specific primers against the target region with *TH* with optimal melting temperatures of 60± 2°C for multiplexed reactions (Supplemental Table 5). We amplified each sample using 10ul reactions with 5X multiplexed primers (10uM), 3X bisulfite converted DNA and 2X KAPA HiFi HotStart Uracil+ ReadyMix (Kapa Biosystems, MA, USA). Each strand was amplified separately, and after two rounds of paramagnetic bead purification at 0.8X, amplicons from both strands were combined and amplified for 10 additional cycles to add custom primer sequences in 20 ul reactions consisting of 2.5X sample, 5X combined CS1 and CS2 primers (10uM) and 2X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, MA, USA). Samples were indexed for 10 cycles in a 20ul reaction consisting of 2.5X amplicons, 5X indexing primers (10uM) and 2X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, MA, USA). Each indexed library went through two rounds of double ended bead purification (final ratio 0.8X) for amplicon size selection (400-700bp). Final library concentrations and quality control was performed on the Agilent 2200 TapeStation (Agilent Technologies, CA, USA) before samples were pooled and sequenced.

MiSeq Sequencing

We pooled libraries to a final concentration of 2nM and included a 5-10% PhiX spike-in control before sequencing on the Illumina MiSeq platform (Illumina, San Diego, CA) using customized 300bp paired end sequencing as described in (Chen et al. 2017). All quality control and read alignment, without removing duplicates, were performed in-house and methylation was calculated as the percent of reads containing cytosine rather than thymidine at each position. After removing statistical outliers (ROUT test; Q=1%), mean methylation between cases and controls was compared using one-tailed Mann-Whitney U or Student's t-tests based on Shapiro Wilks normality which was determined separately for each comparison.

Fluorescence Activated Cell Sorting

Nuclear extraction and labelling

In order to liberate intact nuclei the caudate nucleus and nucleus accumbens (n= 109 total) , we homogenized 50mg of frozen tissue in nuclei buffer containing 10mM PIPES (pH 7.4), 10mM KCl, 2mM $MgCl_2$, 1mM DTT, 0.1% TritonX-100 and 10X Protease Inhibitor Cocktail (Sigma Aldrich, Darmstadt, Germany). Homogenates were passed through a 30% sucrose gradient in nuclei buffer in order to separate nuclei from cellular debris, then after a wash with nuclei buffer, nuclei pellets were resuspended in blocking buffer containing 0.5% bovine serum albumin in 10X normal goat serum. Each sample was co-incubated with the DNA labelling dye DRAQ5 (1:300) (ThermoFisher, Waltham, MA) and an anti-NeuN-PE antibody (1:300) (cat no. FCMAB317PE, Millipore, Darmstadt, Germany) for 60 min at room temperature, then passed through 40uM filter caps before sorting.

Nuclei Sorting

Labelled nuclear extracts were processed on our in-house BDFACSAria III platform (BD Biosciences, San Jose, CA) according to technical specifications provided by the company. We used BD FACSDIVA software (BD Biosciences, San Jose, CA) to first isolate single, intact nuclei based on DRAQ5 fluorescence at the 730/45-A filter (DRAQ5), then to sort neuronal from non-neuronal nuclei based on fluorescence detected by the 585/42 filter (PE). Sorted nuclear fractions were stored at -20°C until DNA extraction.

Nuclear DNA Extraction

We incubated nuclear fractions with 50X protease (Qiagen, Montreal, Canada) at 56°C for at least 12 hours to ensure thorough digestion of the nuclear membranes. Liberated DNA was precipitated onto 0.2X Agencourt AMPure XP (Beckman Coulter, Brea, CA) beads after adding 20% PEG-8000 2.5M

NaCl to a final PEG concentration of 10%. The beads were washed twice in a magnetic stand with 70% EthOH, then DNA was eluted in MilliQ H_2O. We measured the concentration of each DNA sample using Quant-iT PicoGreen dsDNA assays (ThermoFisher, Waltham, MA) according to manufacturer specifications. 500ng of DNA from each fraction was bisulfite converted using the EpiTect fast bisulfite conversion kit (QIAGEN, Cat# 59824) according to the standard protocol, and targeted methylation amplicon libraries were constructed and sequenced as described above.

Luciferase Enhancer Assay

Primer design and PCR

Primers were designed, using the PrimerQuest tool from IDT (http://www.idtdna.com/primerquest/home/index) to amplify the DNA region Chr11 : 2187899-2188296 (hg19). Primers were chosen based on their melting temperature and their ability not to generate primer dimers. Their specificities were assessed by using the NCBI primer blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and the UCSC genome browser (https://genome.ucsc.edu/). The PCR products were generated by two successive PCRs, each followed by purification with standard columns (Qiagen). A 398bp fragment was obtained using the primer a and b (Supplemental Table 6). Using the first purified PCR product as template, secondary PCR products containing restriction sites were generated with the primers c and d (Supplemental Table 6).

Cloning

The PCR products were cloned into the TOPO 2.1 vector (Life Technologies) and TOP10 bacteria were transformed with the ligation products. Positive clones were selected based on the absence of Beta Galactosidase expression. Plasmids were extracted using the Qiagen miniprep kit according to standard protocols and sanger sequenced at the Genome Quebec Innovation Center using the M13 primers.

The inserts identified in the topo TA cloning vector were released from the vector with a double digestion using AvrII (R0174S) and SpeI (R0133S-New England Biolabs) enzymes and gel purification (Qiagen). pCpGfree promoter lucia vectors (Invitrogen) were digested under the same conditions and were dephosphorylated using Antarctic Phosphatase (M0289S-New England Biolabs) followed by column purification (Qiagen). The inserts and pCpGfree vector were ligated overnight at 16°C using T4 Ligase (M0202M-New England Biolabs). GT115 competent bacteria were transformed with the ligation products and plated on LB agar medium with Zeocin (InVivogen). Isolated colonies were then regrown in liquid medium and plasmids were extracted using the Qiagen miniprep kit (Qiagen). The plasmids were verified by PCR using primer designed on the pCpGfree vector (primers e and f) (Supplemental Table 6) and sanger sequenced with the same primers at the Genome Quebec Innovation Center (Montreal, Canada). Clones containing the fragments of interest were regrown in 200ml of LB medium with Zeocin (Invitrogen). Plasmids were re-extracted using a fast midiprep kit (Qiagen). Finally, the final construct was verified by sequencing.

In vitro methylation

The constructs in pCpGfree-promoter-Lucia vector were methylated *in vitro* by treatment with the Sss1 methylase (M0226L-New England Biolabs) and S-Adenin-Methionine for 6 hours and then purified. The efficiency of the *in vitro* methylation was assessed by HPA2 digestion (R0171S-New England Biolabs) of native and methylated vectors followed by agarose gel electrophoresis.

Cell Culture

HEK293 cells were seeded in 6 well plates for expansion in DMEM medium (InVitroGen) supplemented with 10% FBS and penicillin/streptomycin (Life Technologies). Cells were then split in 24-well plates 24 hours before transfection. The cells were transfected at 50 - 60% of confluence with 70ng

of pCpGfree vector and 70ng of pGl3 control vector using the Jetprime transfection kit (Polyplus). Each condition was treated in five biological replicates.

Luciferase assay

The cell culture medium and the cellular extracts were collected 24 hours after transfection. The luciferase activities were assessed in both fractions using a Dual Luciferase reporter assay Kit (Promega) and a Berthold luminometer. Data were acquired using the Simplicity 4.2 Sofware.

Statistics.

A one-way ANOVA analysis was performed to assess inter group differences for luciferase activities. Post-hoc T test were performed and corrected using Bonferroni method to analyse group differences. All the statistical analyses were done using SPSS 20 software.

qPCR Gene expression assays

For expression analysis, we used extracted RNA from our RNA sequencing experiments. Nucleic acid concentration was determined via nanodrop and RNA integrity numbers (RIN) were as follows (mean ± s.d.): 8.4 ± 0.78 for controls and 7.9 ± 1.3 for cases from the caudate, and 8.1 ± 0.6 for controls and 8.1 ± 1.2 for cases in the nucleus accumbens. Aliquots of 25ng/ul were used to generate standard curves, and diluted 1 in 5 for expression analysis.

The expression of *TH* and *EGR1* was determined using quantitative reverse transcription PCR (RT-qPCR). We used pre-designed probe based assays for *TH* (Hs00165941_m1) and *EGR1* (Hs00152928_m1) with FAM-TAMRA dyes (Taqman assays, Thermo Fisher), and ran 10ul assays on an Applied Biosystems QuantStudio 6 instrument under default cycling conditions (ThermoFisher). Expression was calculated based on standard curves and normalized against *GAPDH* (Hs02758991_g1)

using QuantStudio Real-Time PCR Software. Normalized expression was compared between cases and controls using two-tailed Student's t-tests in GraphPad Prism 6 (www.graphpad.com).

Acknowledgments

We are deeply grateful to the families of the subjects used in this study. We would also like to acknowledge the teams of technicians at the Miami Brain Endowment Bank™ and the past and current members of the McGill Group for Suicide Studies for their assistance and support.

This work was supported by a Canadian Institute of Health Research Doctoral Fellowship awarded to K.V., National Institute of Drug Abuse Grants DA033684 awarded to G.T and D.M., and P01 DA047233 to E.N.

Author Contributions

Manuscript preparation: K.V., Experimental design and data collection: K.V., G.G.C., G.M., L.F., Data Analysis: K.V., C.E., A.B., J-F.T., Animal experiments: B.L, E.C., Resources and support: E.N., C.N., N.M., D.C.M., G.T.

Competing Interests

The authors declare no competing interests.

6. References

Albertson, Dawn N., Barb Pruetz, Carl J. Schmidt, Donald M. Kuhn, Gregory Kapatos, and Michael J. Bannon. 2004. “Gene Expression Profile of the Nucleus Accumbens of Human Cocaine Abusers: Evidence for Dysregulation of Myelin.” *Journal of Neurochemistry* 88.

Baker-Andresen, Danay, Qiongyi Zhao, Xiang Li, Bianca Jupp, Rose Chesworth, Andrew J. Lawrence, and Timothy Bredy. 2015. “Persistent Variations in Neuronal DNA Methylation Following Cocaine Self-Administration and Protracted Abstinence in Mice.” *Neuroepigenetics* 4:1–11.

Balda, Mara A., Karen L. Anderson, and Yossef Itzhak. 2009. “The Neuronal Nitric Oxide Synthase (NNOS) Gene Contributes to the Regulation of Tyrosine Hydroxylase (TH) by Cocaine.” *Neuroscience Letters* 457(3):120–24.

Barrera, Víctor and Miguel A. Peinado. 2012. “Evaluation of Single CpG Sites as Proxies of CpG Island Methylation States at the Genome Scale.” *Nucleic Acids Research* 40(22):11490–98.

Beitner-Johnson, Dana, Xavier Guitart, and Eric J. Nestler. 1991. “Dopaminergic Brain Reward Regions of Lewis and Fischer Rats Display Different Levels of Tyrosine Hydroxylase and Other Morphine- and Cocaine-Regulated Phosphoproteins.” *Brain Research* 561(1):147–50.

Belin, David and Barry Everitt. 2008. “Cocaine Seeking Habits Depend upon Dopamine-Dependent Serial Connectivity Linking the Ventral with the Dorsal Striatum.” *Neuron* 57(3):432–41.

Bernstein, Bradley E., Ewan Birney, Ian Dunham, Eric D. Green, Chris Gunter, and Michael Snyder. 2012. “An Integrated Encyclopedia of DNA Elements in the Human Genome.” *Nature* 489(7414):57–74.

Bhat, R. V., A. J. Cole, and J. M. Baraban. 1992. “Role of Monoamine Systems in Activation of Zif268 by Cocaine.” *Journal of Psychiatry and Neuroscience* 17(3):94–102.

Cavalcante, Raymond G. and Maureen A. Sartor. 2017. “Annotatr: Genomic Regions in Context.” *Bioinformatics* 33(15):2381–83.

Chen, Gary G., Alpha B. Diallo, Raphaël Poujol, Corina Nagy, Alfredo Staffa, Kathryn Vaillancourt, Pierre-Eric Lutz, Vanessa K. Ota, Deborah C. Mash, Gustavo Turecki, and Carl Ernst. 2014. “BisQC: An Operational Pipeline for Multiplexed Bisulfite Sequencing.” *BMC Genomics* 15:290.

Chen, Gary G., Jeffrey A. Gross, Pierre-Eric Lutz, Kathryn Vaillancourt, Gilles Maussion, Alexandre Bramoulle, Jean-François Théroux, Elena S. Gardini, Ulrike Ehlert, Geneviève Bourret, Aurélie Masurel, Pierre Lepage, Naguib Mechawar, Gustavo Turecki, and Carl Ernst. 2017. “Medium Throughput Bisulfite Sequencing for Accurate Detection of 5-Methylcytosine and 5-Hydroxymethylcytosine.” *BMC Genomics* 18(1):96.

Dumitriu, Dani, Quincey Laplant, Yael S. Grossman, Caroline Dias, William G. Janssen, Scott J. Russo, John H. Morrison, and Eric J. Nestler. 2012. “Subregional, Dendritic Compartment, and Spine Subtype Specificity in Cocaine Regulation of Dendritic Spines in the Nucleus Accumbens.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 32(20):6957–66.

Feng, Jian, Matthew Wilkinson, Xiaochuan Liu, Immanuel Purushothaman, Deveroux Ferguson, Vincent Vialou, Ian Maze, Ningyi Shao, Pamela Kennedy, Jawook Koo, Caroline Dias, Benjamin Laitman,

Victoria Stockman, Quincey Laplant, Michael Cahill, Eric J. Nestler, and Li Shen. 2014. “Chronic Cocaine-Regulated Epigenomic Changes in Mouse Nucleus Accumbens.” *Genome Biology* 15(4):R65.

Fonteneau, M., D. Filliol, P. Anglard, K. Befort, P. Romieu, and J. Zwiller. 2017. “Inhibition of DNA Methyltransferases Regulates Cocaine Self-Administration by Rats: A Genome-Wide DNA Methylation Study.” *Genes, Brain and Behavior* 16(3):313–27.

Fritz, Michael, Rana El Rawas, Ahmad Salti, Sabine Klement, Michael T. Bardo, Georg Kemmler, Georg Dechant, Alois Saria, and Gerald Zernig. 2011. “Reversal of Cocaine-Conditioned Place Preference and Mesocorticolimbic Zif268 Expression by Social Interaction in Rats.” *Addiction Biology* 16(2):273–84.

Gandal, Michael J., Pan Zhang, Evi Hadjimichael, Rebecca L. Walker, Chao Chen, Shuang Liu, Hyejung Won, Harm van Bakel, Merina Varghese, Yongjun Wang, Annie W. Shieh, Jillian Haney, Sepideh Parhami, Judson Belmont, Minsoo Kim, Patricia Moran Losada, Zenab Khan, Justyna Mleczko, Yan Xia, Rujia Dai, Daifeng Wang, Yucheng T. Yang, Min Xu, Kenneth Fish, Patrick R. Hof, Jonathan Warrell, Dominic Fitzgerald, Kevin White, Andrew E. Jaffe, Mette A. Peters, Mark Gerstein, Chunyu Liu, Lilia M. Iakoucheva, Dalila Pinto, Daniel H. Geschwind, and Daniel H. Geschwind. 2018. “Transcriptome-Wide Isoform-Level Dysregulation in ASD, Schizophrenia, and Bipolar Disorder.” *Science* 362(6420):eaat8127.

Gervasi, Noreen M., Shane S. Scott, Armaz Aschrafi, Jenna Gale, Sanah N. Vohra, Margaret A. Macgibeny, Amar N. Kar, Anthony E. Gioio, and Barry B. Kaplan. 2016. “The Local Expression and Trafficking of Tyrosine Hydroxylase MRNA in the Axons of Sympathetic Neurons.” *RNA* 22(6):883–95.

Grimm, J. W., Y. Shaham, and B. T. Hope. 2002. “Effect of Cocaine and Sucrose Withdrawal Period on Extinction Behavior, Cue-Induced Reinstatement, and Protein Levels of the Dopamine Transporter and Tyrosine Hydroxylase in Limbic and Cortical Areas in Rats.” *Behavioural Pharmacology* 13(5–6):379–88.

GTEx Consortium, The GTEx. 2015. “Human Genomics. The Genotype-Tissue Expression (GTEx) Pilot Analysis: Multitissue Gene Regulation in Humans.” *Science (New York, N.Y.)* 348(6235):648–60.

Hámor, Peter U., Mariola J. Edelmann, Christina Gobin, and Marek Schwendt. 2020. “Long-Term Changes in the Central Amygdala Proteome in Rats with a History of Chronic Cocaine Self-Administration.” *Neuroscience* 443:93–109.

Hannon, Eilis, Sarah J. Marzi, Leonard S. Schalkwyk, and Jonathan Mill. 2019. “Genetic Risk Variants for Brain Disorders Are Enriched in Cortical H3K27ac Domains.” *Molecular Brain* 12(1):7.

Hearing, M. C., R. E. See, and J. F. McGinty. 2008. “Relapse to Cocaine-Seeking Increases Activity-Regulated Gene Expression Differentially in the Striatum and Cerebral Cortex of Rats Following Short or Long Periods of Abstinence.” *Brain Structure and Function* 213(1–2):215–27.

Ibáñez-Sandoval, Osvaldo, Fatuel Tecuapetla, Bengi Unal, Fulva Shah, Tibor Koós, and James M. Tepper. 2010. “Electrophysiological and Morphological Characteristics and Synaptic Connectivity of Tyrosine Hydroxylase-Expressing Neurons in Adult Mouse Striatum.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 30(20):6999–7016.

Kaminer, Jaime, Diego Espinoza, Shaznaan Bhimani, James M. Tepper, Tibor Koos, and Michael W.

Shiflett. 2019. “Loss of Striatal Tyrosine-Hydroxylase Interneurons Impairs Instrumental Goal-Directed Behavior.” *European Journal of Neuroscience* 50(4):2653–62.

Kozlenkov, Alexey, Minghui Wang, Panos Roussos, Sergei Rudchenko, Mihaela Barbu, Marina Bibikova, Brandy Klotzle, Andrew J. Dwork, Bin Zhang, Yasmin L. Hurd, Eugene V. Koonin, Michael Wegner, and Stella Dracheva. 2015. “Substantial DNA Methylation Differences between Two Major Neuronal Subtypes in Human Brain.” *Nucleic Acids Research* 44(6):2593–2612.

Kundaje, Anshul, Wouter Meuleman, Jason Ernst, Misha Bilenky, Angela Yen, Alireza Heravi-Moussavi, Pouya Kheradpour, Zhizhuo Zhang, Jianrong Wang, Michael J. Ziller, Viren Amin, John W. Whitaker, Matthew D. Schultz, Lucas D. Ward, Abhishek Sarkar, Gerald Quon, Richard S. Sandstrom, Matthew L. Eaton, Yi-Chieh Wu, Andreas R. Pfenning, Xinchen Wang, Melina Claussnitzer, Yaping Liu, Cristian Coarfa, R. Alan Harris, Noam Shoresh, Charles B. Epstein, Elizabeta Gjoneska, Danny Leung, Wei Xie, R. David Hawkins, Ryan Lister, Chibo Hong, Philippe Gascard, Andrew J. Mungall, Richard Moore, Eric Chuah, Angela Tam, Theresa K. Canfield, R. Scott Hansen, Rajinder Kaul, Peter J. Sabo, Mukul S. Bansal, Annaick Carles, Jesse R. Dixon, Kai-How Farh, Soheil Feizi, Rosa Karlic, Ah-Ram Kim, Ashwinikumar Kulkarni, Daofeng Li, Rebecca Lowdon, GiNell Elliott, Tim R. Mercer, Shane J. Neph, Vitor Onuchic, Paz Polak, Nisha Rajagopal, Pradipta Ray, Richard C. Sallari, Kyle T. Siebenthall, Nicholas A. Sinnott-Armstrong, Michael Stevens, Robert E. Thurman, Jie Wu, Bo Zhang, Xin Zhou, Arthur E. Beaudet, Laurie A. Boyer, Philip L. De Jager, Peggy J. Farnham, Susan J. Fisher, David Haussler, Steven J. M. Jones, Wei Li, Marco A. Marra, Michael T. McManus, Shamil Sunyaev, James A. Thomson, Thea D. Tlsty, Li-Huei Tsai, Wei Wang, Robert A. Waterland, Michael Q. Zhang, Lisa H. Chadwick, Bradley E. Bernstein, Joseph F. Costello, Joseph R. Ecker, Martin Hirst, Alexander Meissner, Aleksandar Milosavljevic, Bing Ren, John A. Stamatoyannopoulos, Ting Wang, Manolis Kellis, and Manolis Kellis. 2015. “Integrative Analysis of 111 Reference Human Epigenomes.” *Nature* 518(7539):317–30.

LaPlant, Quincey, Vincent Vialou, Herbert E. Covington, Dani Dumitriu, Jian Feng, Brandon L. Warren, Ian Maze, David M. Dietz, Emily L. Watts, Sergio D. Iñiguez, Ja Wook Koo, Ezekiell Mouzon, William Renthal, Fiona Hollis, Hui Wang, Michele A. Noonan, Yanhua Ren, Amelia J. Eisch, Carlos A. Bolaños, Mohamed Kabbaj, Guanghua Xiao, Rachael L. Neve, Yasmin L. Hurd, Ronald S. Oosting, Gouping Fan, John H. Morrison, and Eric J. Nestler. 2010. “Dnmt3a Regulates Emotional Behavior and Spine Plasticity in the Nucleus Accumbens.” *Nature Neuroscience* 13(9):1137–43.

Letchworth, S. R., M. a Nader, H. R. Smith, D. P. Friedman, and L. J. Porrino. 2001. “Progression of Changes in Dopamine Transporter Binding Site Density as a Result of Cocaine Self-Administration in Rhesus Monkeys.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 21(8):2799–2807.

Lister, Ryan, Eran a Mukamel, Joseph R. Nery, Mark Urich, Clare a Puddifoot, Nicholas D. Johnson, Jacinta Lucero, Yun Huang, Andrew J. Dwork, Matthew D. Schultz, Miao Yu, Julian Tonti-Filippini, Holger Heyn, Shijun Hu, Joseph C. Wu, Anjana Rao, Manel Esteller, Chuan He, Fatemeh G. Haghighi, Terrence J. Sejnowski, M. Margarita Behrens, and Joseph R. Ecker. 2013. “Global Epigenomic Reconfiguration during Mammalian Brain Development.” *Science (New York, N.Y.)* 341(6146):1237905.

Lobo, Mary Kay, Stanislav L. Karsten, Michelle Gray, Daniel H. Geschwind, and X. William Yang. 2006. “FACS-Array Profiling of Striatal Projection Neuron Subtypes in Juvenile and Adult Mouse Brains.” *Nature Neuroscience* 9(3):443–52.

Logan, Ryan W., Puja K. Parekh, Gabrielle N. Kaplan, Darius D. Becker-Krail, Wilbur P. Williams, Shintaro

Yamaguchi, Jun Yoshino, Micah A. Shelton, Xiyu Zhu, Hui Zhang, Spencer Waplinger, Ethan Fitzgerald, Jeffrey Oliver-Smith, Poornima Sundarvelu, John F. Enwright, Yanhua H. Huang, and Colleen A. McClung. 2019. "NAD+ Cellular Redox and SIRT1 Regulate the Diurnal Rhythms of Tyrosine Hydroxylase and Conditioned Cocaine Reward." *Molecular Psychiatry* 24(11):1668–84.

Luo, Chongyuan, Christopher L. Keown, Laurie Kurihara, Jingtian Zhou, Yupeng He, Junhao Li, Rosa Castanon, Jacinta Lucero, Joseph R. Nery, Justin P. Sandoval, Brian Bui, Terrence J. Sejnowski, Timothy T. Harkins, Eran A. Mukamel, M. Margarita Behrens, and Joseph R. Ecker. 2017. "Single-Cell Methylomes Identify Neuronal Subtypes and Regulatory Elements in Mammalian Cortex." *Science* 357(6351):600–604.

Massart, Renaud, Royi Barnea, Yahav Dikshtein, Matthew Suderman, Oren Meir, Michael Hallett, Pamela Kennedy, Eric J. Nestler, Moshe Szyf, and Gal Yadid. 2015. "Role of DNA Methylation in the Nucleus Accumbens in Incubation of Cocaine Craving." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 35(21):8042–58.

Masserano, Joseph M., Ivory Baker, Naoki Natsukari, and Richard Jed Wyatt. 1996. "Chronic Cocaine Administration Increases Tyrosine Hydroxylase Activity in the Ventral Tegmental Area through Glutaminergic- and Dopaminergic D2-Receptor Mechanisms." *Neuroscience Letters* 217(2–3):73–76.

Maunakea, Alika K., Iouri Chepelev, Kairong Cui, and Keji Zhao. 2013. "Intragenic DNA Methylation Modulates Alternative Splicing by Recruiting MeCP2 to Promote Exon Recognition." *Cell Research* 23(11):1256–69.

Maunakea, Alika K., Raman P. Nagarajan, Mikhail Bilenky, Tracy J. Ballinger, Cletus D'Souza, Shaun D. Fouse, Brett E. Johnson, Chibo Hong, Cydney Nielsen, Yongjun Zhao, Gustavo Turecki, Allen Delaney, Richard Varhol, Nina Thiessen, Ksenya Shchors, Vivi M. Heine, David H. Rowitch, Xiaoyun Xing, Chris Fiore, Maximiliaan Schillebeeckx, Steven J. M. Jones, David Haussler, Marco A. Marra, Martin Hirst, Ting Wang, and Joseph F. Costello. 2010. "Conserved Role of Intragenic DNA Methylation in Regulating Alternative Promoters." *Nature* 466(7303):253–57.

Moratalla, R., H. A. Robertson, and A. M. Graybiel. 1992. "Dynamic Regulation of NGFI-A (Zif268, Egr1) Gene Expression in the Striatum." *Journal of Neuroscience* 12(7):2609–22.

Nestler, Eric J. and Christian Lüscher. 2019. "The Molecular Basis of Drug Addiction: Linking Epigenetic to Synaptic and Circuit Mechanisms." *Neuron* 102(1):48–59.

Saunders, Arpiar, Evan Z. Macosko, Alec Wysoker, Melissa Goldman, Fenna M. Krienen, Heather de Rivera, Elizabeth Bien, Matthew Baum, Laura Bortolin, Shuyu Wang, Aleksandrina Goeva, James Nemesh, Nolan Kamitaki, Sara Brumbaugh, David Kulp, and Steven A. McCarroll. 2018. "Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain." *Cell* 174(4):1015-1030.e16.

Savell, Katherine E., Jennifer J. Tuscher, Morgan E. Zipperly, Corey G. Duke, Robert A. Phillips, Allison J. Bauman, Saakshi Thukral, Faraz A. Sultan, Nicholas A. Goska, Lara Ianov, and Jeremy J. Day. 2020. "A Dopamine-Induced Gene Expression Signature Regulates Neuronal Function and Cocaine Response." *Science Advances* 6(26).

Schmidt, E. F., M. A. Sutton, C. A. Schad, D. A. Karanian, E. S. Brodkin, and D. W. Self. 2001. "Extinction Training Regulates Tyrosine Hydroxylase during Withdrawal from Cocaine Self-Administration." *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 21(7).

Schoenfelder, Stefan and Peter Fraser. 2019. “Long-Range Enhancer–Promoter Contacts in Gene Expression Control.” *Nature Reviews Genetics* 20(8):437–55.

Sheffield, Nathan C. and Christoph Bock. 2016. “LOLA: Enrichment Analysis for Genomic Region Sets and Regulatory Elements in R and Bioconductor.” *Bioinformatics* 32(4):587–89.

Thomas, Paul D., Michael J. Campbell, Anish Kejariwal, Huaiyu Mi, Brian Karlak, Robin Daverman, Karen Diemer, Anushya Muruganujan, and Apurva Narechania. 2003. “PANTHER: A Library of Protein Families and Subfamilies Indexed by Function.” *Genome Research* 13(9):2129–41.

Thomsen, Morgane and S. Barak Caine. 2007. “Intravenous Drug Self-Administration in Mice: Practical Considerations.” *Behavior Genetics* 37(1):101–18.

Tian, Weiping, Mei Zhao, Min Li, Tianbao Song, Min Zhang, Li Quan, Shengbin Li, and Zhong Sheng Sun. 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.

Todtenkopf, M. S., K. R. De Leon, and J. R. Stellar. 2000. “Repeated Cocaine Treatment Alters Tyrosine Hydroxylase in the Rat Nucleus Accumbens.” *Brain Research Bulletin* 52(5):407–11.

Tomasi, Dardo, Nora D. Volkow, Ruiliang Wang, Jean H. Carrillo, Thomas Maloney, Nelly Alia-Klein, Patricia a Woicik, Frank Telang, and Rita Z. Goldstein. 2010. “Disrupted Functional Connectivity with Dopaminergic Midbrain in Cocaine Abusers.” *PloS One* 5(5):e10815.

Vaillancourt, Kathryn, Carl Ernst, Deborah Mash, and Gustavo Turecki. 2017. “DNA Methylation Dynamics and Cocaine in the Brain: Progress and Prospects.” *Genes* 8(5):138.

Vaillancourt, Kathryn, Jennie Yang, Gary G. Chen, Volodymyr Yerko, Jean François Théroux, Zahia Aouabed, Alberto Lopez, Kimberly C. Thibeault, Erin S. Calipari, Benoit Labonté, Naguib Mechawar, Carl Ernst, Corina Nagy, Thierry Forné, Eric J. Nestler, Deborah C. Mash, and Gustavo Turecki. 2020. “Cocaine-Related DNA Methylation in Caudate Neurons Alters 3D Chromatin Structure of the IRXA Gene Cluster.” *Molecular Psychiatry*.

Valjent, Emmanuel, Benjamin Aubier, Anne Gaëlle Corbillé, Karen Brami-Cherrier, Piotr Topilko, Jean Antoine Girault, and Denis Hervé. 2006. “Plasticity-Associated Gene Krox24/Zif268 Is Required for Long-Lasting Behavioral Effects of Cocaine.” *Journal of Neuroscience* 26(18):4956–60.

Volkow, N. D., G. J. Wang, J. S. Fowler, J. Logan, S. J. Gatley, A. Gifford, R. Hitzemann, Y. S. Ding, and N. Pappas. 1999. “Prediction of Reinforcing Responses to Psychostimulants in Humans by Brain Dopamine D2 Receptor Levels.” *The American Journal of Psychiatry* 156(9):1440–43.

Volkow, N. D., G. J. Wang, F. Telang, J. S. Fowler, J. Logan, A. R. Childress, M. Jayne, Y. Ma, and C. Wong. 2006. “Cocaine Cues and Dopamine in Dorsal Striatum: Mechanism of Craving in Cocaine Addiction.” *Journal of Neuroscience* 26(24):6583–88.

Walker, Deena M., Hannah M. Cates, Yong Hwee E. Loh, Immanuel Purushothaman, Aarthi Ramakrishnan, Kelly M. Cahill, Casey K. Lardner, Arthur Godino, Hope G. Kronman, Jacqui Rabkin, Zachary S. Lorsch, Philipp Mews, Marie A. Doyle, Jian Feng, Benoit Labonté, Ja Wook Koo, Rosemary C. Bagot, Ryan W. Logan, Marianne L. Seney, Erin S. Calipari, Li Shen, and Eric J. Nestler. 2018. “Cocaine Self-Administration Alters Transcriptome-Wide Responses in the Brain’s Reward Circuitry.” *Biological Psychiatry* 84(12):867–80.

Wright, Katherine N., Fiona Hollis, Florian Duclot, Amanda M. Dossat, Caroline E. Strong, T. Chase

Francis, Roger Mercer, Jian Feng, David M. Dietz, Mary Kay Lobo, Eric J. Nestler, and Mohamed Kabbaj. 2015. “Methyl Supplementation Attenuates Cocaine-Seeking Behaviors and Cocaine-Induced c-Fos Activation in a DNA Methylation-Dependent Manner.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 35(23):8948–58.

Xenias, Harry S., Osvaldo Ibáñez-Sandoval, Tibor Koós, and James M. Tepper. 2015. “Are Striatal Tyrosine Hydroxylase Interneurons Dopaminergic?” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 35(16):6584–99.

Zandarashvili, Levani, Mark A. White, Alexandre Esadze, and Junji Iwahara. 2015. “Structural Impact of Complete CpG Methylation within Target DNA on Specific Complex Formation of the Inducible Transcription Factor Egr-1.” *FEBS Letters* 589(15):1748–53.

7. Figures

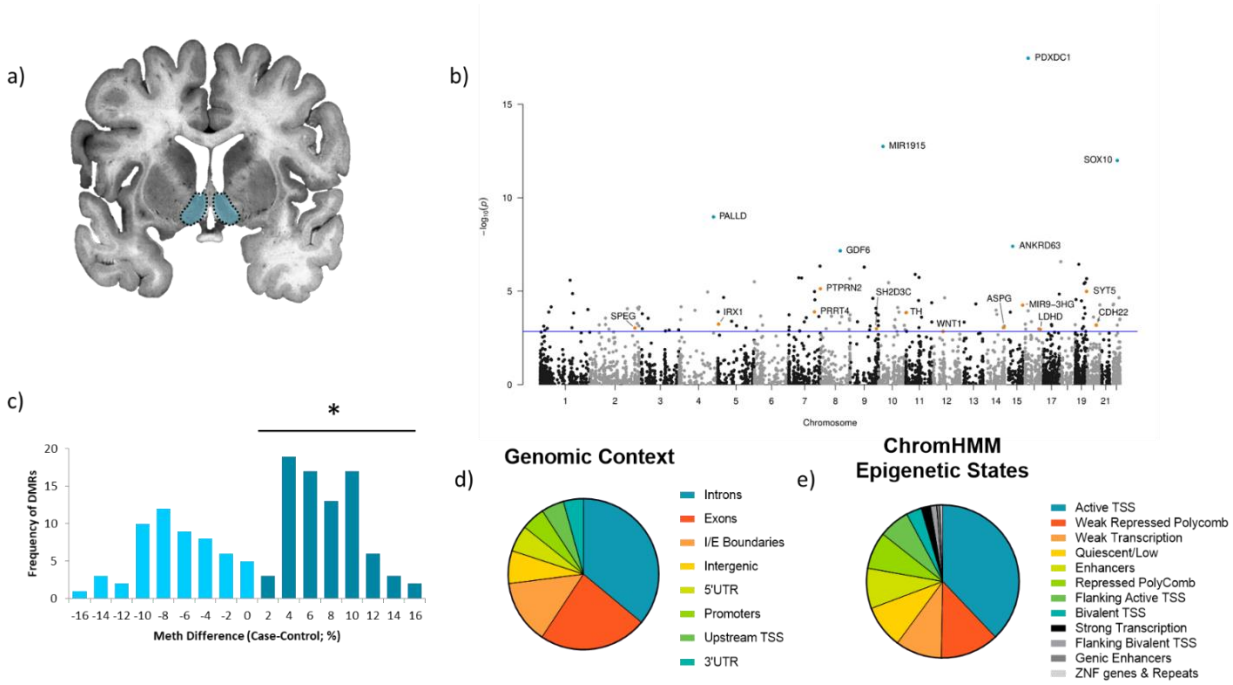


Figure 1. Cocaine dependence is associated with widespread alterations in DNA methylation in the human nucleus accumbens. a) Nucleus accumbens tissue from 25 individuals with a history of cocaine dependence and 25 drug-free controls was used to generate RRBS methylation data. b) Differentially methylated regions (DMRs) were distributed across autosomes, with the most significantly different regions indicated by blue dots, and those which overlapped with data from the caudate nucleus in orange. c) DMRs are significantly skewed towards hypermethylation, d) are found primarily within gene bodies, e) and are enriched in loci that overlap with active TSSs. TSS= transcription start site, UTR=untranslated region, I/E = intron/exon, ZNF = zinc finger, * $p < 0.05$

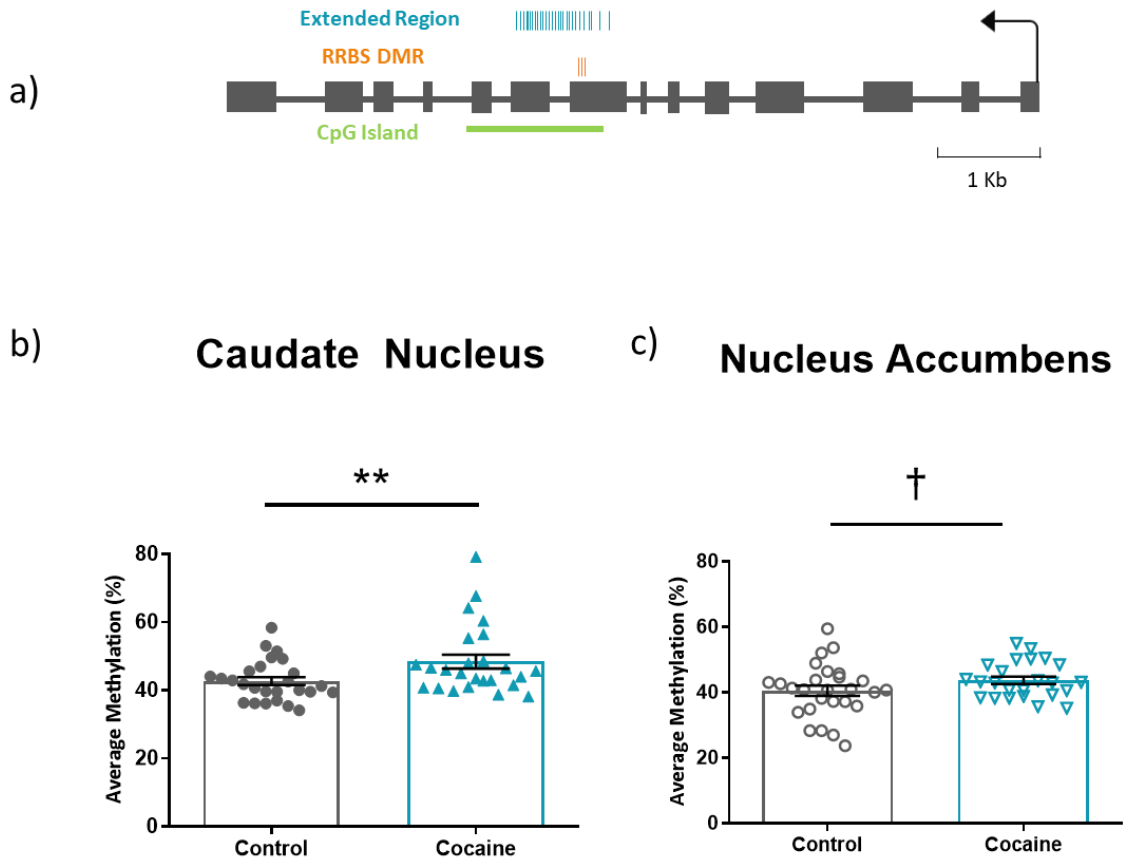


Figure 2. Exon 8/9 of TH contains a CpG island that is hypermethylated in the cocaine group. a) Schematic diagram of the tyrosine hydroxylase gene containing the position of CpGs identified in the methylome wide analysis (orange, RRBS DMR), as well as those which were sequenced in follow-up experiments (blue, Extended Region). b) Average methylation of the extended region is significantly higher in the cocaine group (n=25) compared to controls (n=26) in the caudate nucleus, and c) there is a trend towards higher methylation in the cocaine group (n=24) compared to controls (n=27) in the nucleus accumbens. Data represented as mean ± s.e.m., ** $p < 0.01$, † $p = 0.058$

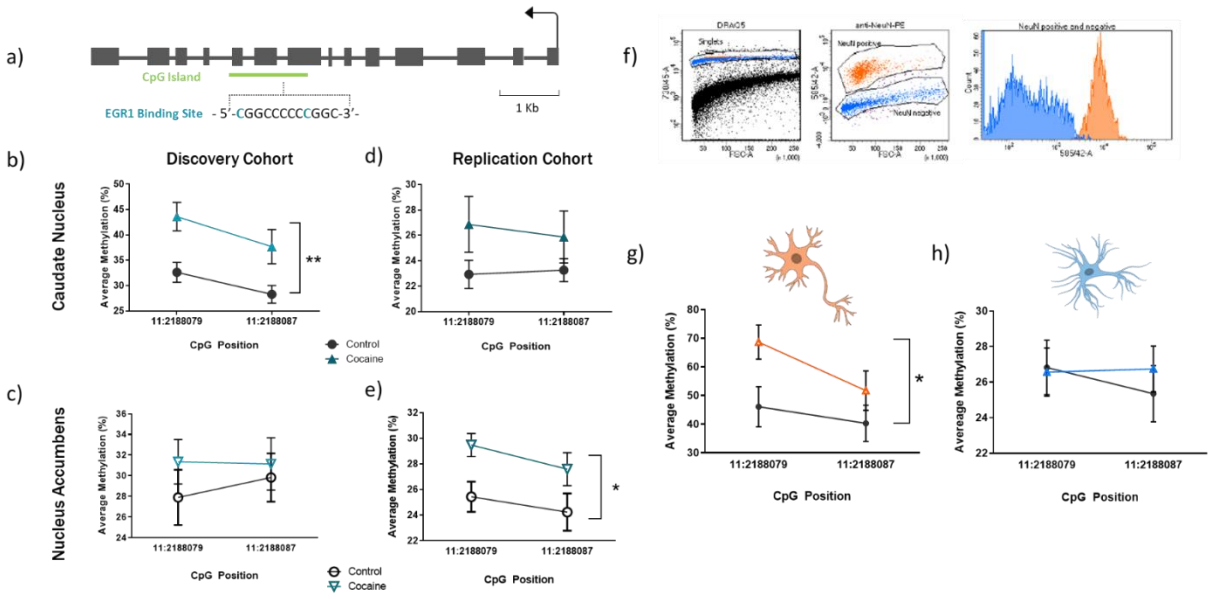


Figure 3. Striatal hypermethylation of CpGs within an EGR1 binding site is specific to neurons in the caudate nucleus. a) The intragenic CpG island in TH contains a potential EGR1 binding site containing 2 flanking CpGs that are b) hypermethylated within the caudate nucleus of the discovery cohort and e) the nucleus accumbens of the replication cohort (n=17 cases and 18 controls). c) There is a non-significant trend towards hypermethylation within the nucleus accumbens of the discovery cohort and d) caudate nucleus of the replication cohort (n=16 cases and 18 controls). f) Single nuclei from discovery accumbens and caudate were separated using FANS, based on DRAQ5 immunofluorescence, and neuronal nuclei were separated based on NeuN. g) Hypermethylation of the CpGs within the EGR1 binding site is specific to neuronal nuclei (n=22 cases and 20 controls) and not h) non-neuronal nuclei. Data represented as mean ± s.e.m., ** $p < 0.01$, * $p < 0.05$

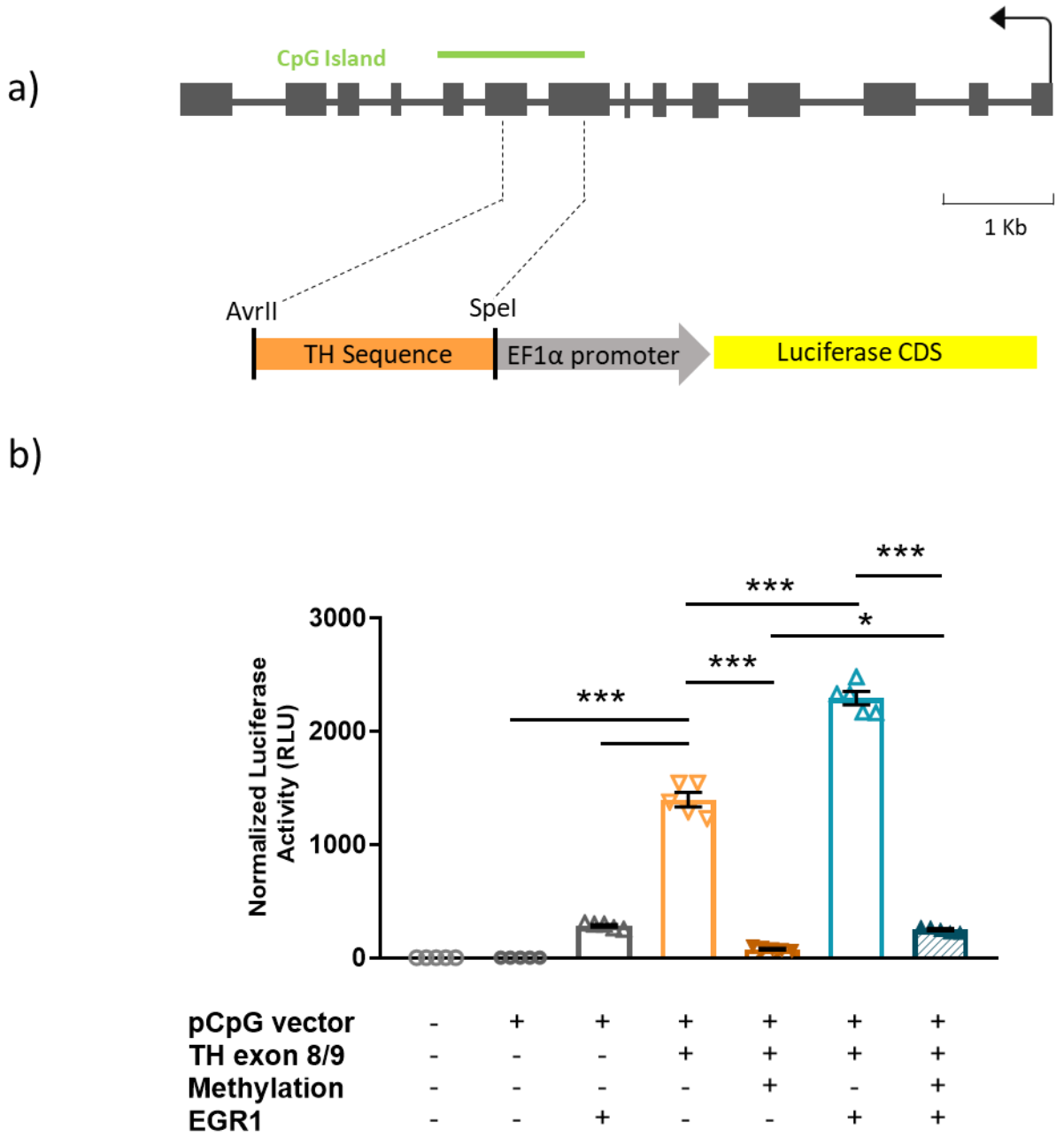


Figure 4. The exon 8/9 region of *TH* functions as an enhancer in vitro and is modulated by methylation and EGR1 overexpression. a) The region containing the cocaine DMR was cloned into a Luciferase enhancer plasmid and transfected into HEK293 cells for functional analyses. b) The target region acts as an enhancer to significantly increase luciferase fluorescence, compared to mock transfected and empty vector controls. Enhancer activity is abolished by DNA methylation and enhanced by co-transfection with an EGR1 overexpression vector (n=5 biological replicates per condition). Data represented as mean ± s.e.m., *** $p < 0.005$, * $p < 0.05$

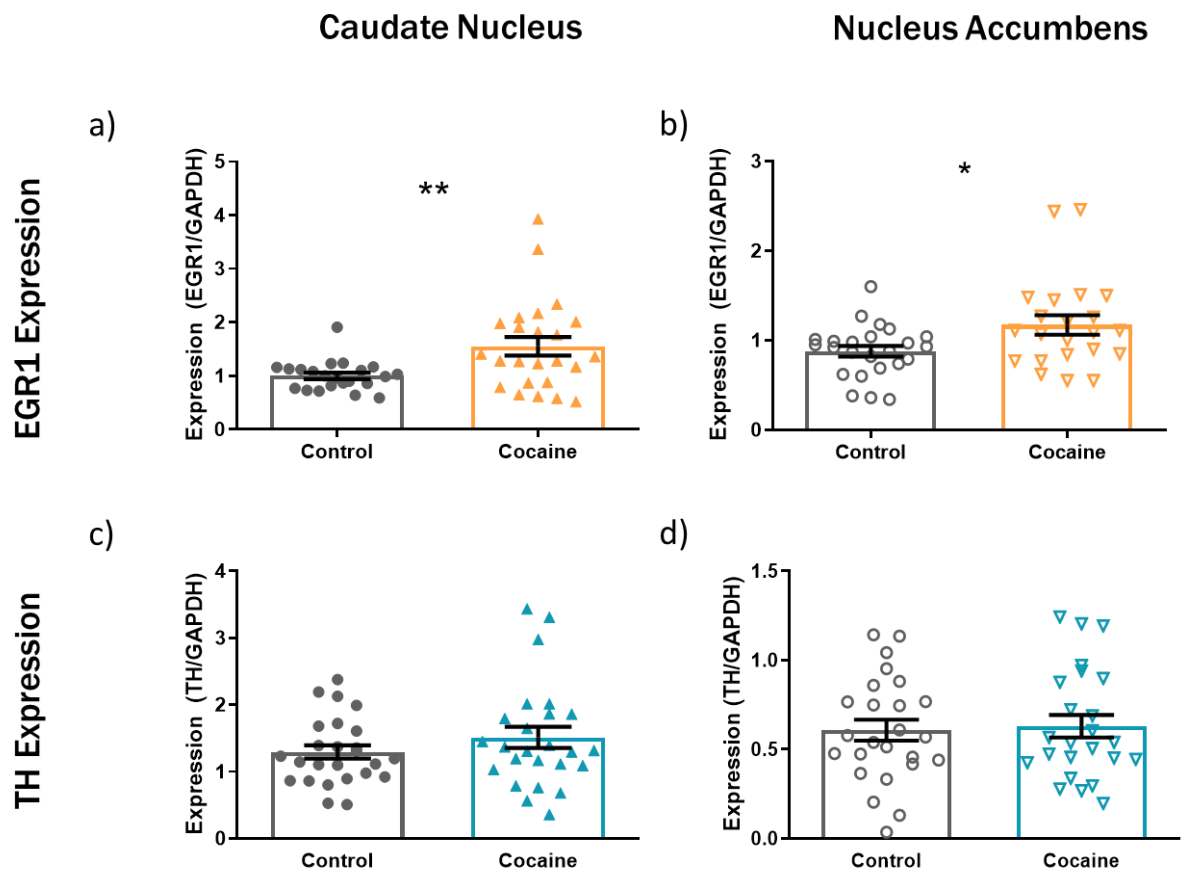


Figure 5. EGR1 expression is increased in the cocaine group, while TH expression remains unchanged. a) There is significantly more EGR1 mRNA in the caudate nucleus and b) nucleus accumbens in the cocaine group (n=24-25) compared to controls (n=25). c) and d) There was no significant difference in TH mRNA expression in either brain region (n=24-25 cases and 25 controls). Data represented as mean ± s.e.m., ** $p < 0.01$, * $p < 0.05$

8. Tables

Table 1. Top hypermethylated region in the nucleus accumbens

Chromosome	From	To	Nearest Gene	Methylation Difference (%)
10	21789249	21789295	MIR1915	15.93
1	147718180	147718225	NBPF8	15.22
11	3244384	3244429	MRGPRE	14.33
19	39798143	39798241	LRFN1	13.41
19	55944580	55944625	SHISA7	13.17
19	44038525	44038544	ZNF575	12.28
13	79183633	79183659	RNF219	11.85
19	55685105	55685146	SYT5	11.14
9	138148440	138148475	LOC401557	10.88
7	149489530	149489549	SSPO	10.50
2	240241219	240241263	HDAC4	10.40
10	91294442	91294456	SLC16A12	10.11
2	74875158	74875201	M1AP	9.82
7	128556152	128556192	KCP	9.71
4	169798988	169799036	PALLD	9.46
5	28927814	28927856	LSP1P3	9.26
11	2188129	2188165	TH	9.17
15	89922099	89922217	MIR9-3HG	9.16
11	65352989	65353032	EHBP1L1	8.79
7	2660963	2661007	IQCE	8.77

Table 2. Differentially methylated regions and genes in common between the nucleus accumbens and the caudate nucleus

Chromosome	From	To	Gene	Nucleus Accumbens Methylation Difference (%)	Caudate Nucleus Methylation Difference (%)
Common DMRs					
2	220313302	220313338	SPEG	-11.5	8.91
11	2188129	2188165	TH	9.17	9.09
16	15083748	15083904	PDXDC1	6.02	3.07
16	75148495	75148543	LDHD	4.32	-6.65
19	55685105	55685146	SYT5	11.1	3.59
Common Genes					
5	3599609	3599704	IRX1		10.7
	3605955	3605982		8.56	
7	157484891	157484931	PTPRN2		-3.61
	157478003	157478040		-8.40	
7	127991684	127991788	PRRT4		-1.72
	127991615	127991788		4.75	
9	130517323	130517347	SH2D3C		5.99
	130517429	130517456		4.82	
12	49373486	49373526	WNT1		6.85
	49374220	49374248		-10.01	
14	104561244	104561286	ASPG		6.29
	104552290	104552330		-9.54	
15	89921805	89921837	MIR9-3HG		9.22
	89922099	89922217		9.16	
20	44879878	44879903	CDH22		9.00
	44803094	44803290		2.84	

9. Supplemental Tables and Figures

Tables

Supplemental Table 1. Demographic information for the discovery cohort

	Controls	Cases
N	25	25
Ethnicity*	11 (A); 7 (C); 11 (H)	13 (A); 6 (C); 10 (H)
Age (years)	38.4 ± 12.5	35.2 ± 8.5
PMI (hours)	15.1 ± 7.2	14.5 ± 7.7
pH	6.4 ± 0.3	6.5 ± 0.3
Sex**	25 (M)	25 (M)

*A= African American; C= Caucasian; H= Hispanic

** M=male

Supplemental Table 2. Reduced Representation Bisulfite Sequencing Statistics

	Average Reads per Library	Alignment (%)	Conversion Efficiency (% ± s.d.)	CpG Coverage at 1X (#)	CpG Coverage at 5X (#)	CpG Coverage at 10X (#)
Cocaine	38331982	57.9	99.0 ± 0.21	2256462	1608983	1258383
Control	28346594	58.4	99.1 ± 0.18	2149014	1464762	1101730

Supplemental Table 3. Differentially methylated region in the nucleus accumbens

1	17475810	17475857	PADI2	3.58
1	22241594	22241616	HSPG2	-14.07
1	34632088	34632123	C1orf94	4.64
1	41982018	41982063	HIVEP3	3.97
1	53791689	53791712	NRD1	6.92
1	147718180	147718225	NBPF8	15.22
1	156878491	156878518	PEAR1	-12.82
1	160053850	160053892	KCNJ9	-5.71
1	165345926	165345957	LMX1A	6.64
1	236157087	236157128	NID1	-8.81
2	43295293	43295339	ZFP36L2	6.90
2	74663630	74663678	RTKN	-10.13
2	74875158	74875201	M1AP	9.82
2	91777521	91777570	AC116050.1	-4.86
2	182521242	182521266	CERKL	-8.08
2	220313302	220313338	SPEG	-11.48
2	231855693	231855741	LOC348761	-8.49
2	234370049	234370089	DGKD	-10.67
2	238472683	238472717	PRLH	-10.02
2	240241219	240241263	HDAC4	10.40
2	241395420	241395464	MIR149	-6.61
3	13696946	13697057	LINC00620	0.88
3	13760173	13760195	LINC00620	-2.99
3	42947590	42947617	ZNF662	2.72
3	128765031	128765196	EFCC1	-2.53
3	147142205	147142339	ZIC1	1.76
3	194117737	194117778	GP5	7.62
4	6660521	6660564	MRFAP1	2.52
4	8863494	8863539	HMX1	8.10
4	77819560	77819600	SOWAHB	6.82
4	140587179	140587217	MGST2	-5.21
4	169798988	169799036	PALLD	9.46
5	1551645	1551687	SDHAP3	-11.88
5	3605955	3605982	IRX1	8.56
5	28927814	28927856	LSP1P3	9.26
5	68711077	68711125	MARVELD2	-4.64
5	92924314	92924354	NR2F1	-9.21
5	139526119	139526191	IGIP	4.19
6	292576	292621	DUSP22	7.32
6	13860679	13860722	MCUR1	-6.28
6	32064082	32064147	TNXB	-4.15
6	41340705	41340752	NCR2	5.52
7	1408565	1408605	MICALL2	5.17

7	1789642	1789675	ELFN1	-14.15
7	2660963	2661007	IQCE	8.77
7	50132821	50132861	ZPBP	6.73
7	54725222	54725255	VSTM2A	5.75
7	64349584	64349632	ZNF273	8.66
7	77648600	77648643	MAGI2	3.41
7	87505884	87505928	DBF4	-12.53
7	127991615	127991788	PRRT4	4.75
7	128556152	128556192	KCP	9.71
7	130132063	130132111	MEST	4.41
7	149489530	149489549	SSPO	10.50
7	155595910	155595957	SHH	4.48
7	157478003	157478040	PTPRN2	-8.40
8	27183146	27183238	PTK2B	3.66
8	53322477	53322510	ST18	-4.68
8	94508105	94508228	LINC00535	2.06
8	97157899	97158021	GDF6	5.67
8	103750914	103750959	KLF10	-10.47
8	140875138	140875160	TRAPPC9	-8.15
8	142528795	142528842	MROH5	3.06
8	143807803	143807844	THEM6	-8.37
8	144361428	144361457	MINCR	-14.62
8	145755735	145755771	ARHGAP39	3.04
9	36137582	36137613	GLIPR2	8.75
9	69785442	69785482	-	-11.34
9	96674623	96674671	BARX1	8.75
9	99482243	99482280	-	-6.67
9	112811535	112811575	AKAP2	2.69
9	128171086	128171126	-	-9.26
9	129245031	129245074	MVB12B	3.50
9	130517429	130517456	SH2D3C	4.82
9	138148440	138148475	LOC401557	10.88
9	138637608	138637630	KCNT1	-10.10
9	139494354	139494401	NOTCH1	-9.95
9	140062475	140062520	LRRC26	2.48
10	21789249	21789295	MIR1915	15.93
10	49674588	49674634	ARHGAP22	3.24
10	77157142	77157181	ZNF503	3.89
10	81539193	81539241	NUTM2B-AS1	-7.72
10	82214986	82215078	TSPAN14	-0.05
10	91294442	91294456	SLC16A12	10.11
10	94822004	94822041	CYP26C1	-1.05
10	101281945	101281982	LINC01475	-9.90
10	127584530	127584601	FANK1	6.59
10	131694687	131694770	EBF3	5.96

10	135278763	135278799	SCART1	-6.97
11	2188129	2188165	TH	9.17
11	3244384	3244429	MRGPRE	14.33
11	46369240	46369288	DGKZ	5.22
11	63529951	63529995	C11orf95	-8.50
11	64901047	64901084	MIR6751	-6.54
11	65352989	65353032	EHBP1L1	8.79
11	77907332	77907379	USP35	-6.86
11	128321703	128321750	ETS1	-12.72
11	128694359	128694406	FLI1	-1.32
12	49374220	49374248	WNT1	-10.01
12	58131090	58131133	AGAP2	-9.10
13	21296167	21296211	IL17D	5.28
13	79183633	79183659	RNF219	11.85
14	65016854	65016901	PPP1R36	-1.17
14	69341470	69341518	ACTN1	8.01
14	100204045	100204074	EML1	7.23
14	104552290	104552330	ASPG	-9.54
14	105635610	105635653	NUDT14	-3.41
15	28342366	28342413	OCA2	3.36
15	40574234	40574253	ANKRD63	3.40
15	89922099	89922217	MIR9-3HG	9.16
16	1104820	1104835	SSTR5-AS	-7.13
16	3070420	3070465	TNFRSF12A	2.16
16	15083748	15083904	PDXDC1	6.02
16	15083985	15084032	PDXDC1	6.36
16	30960806	30960849	ORAI3	-3.54
16	67918743	67918758	NRN1L	1.83
16	75148495	75148543	LDHD	4.32
17	40821978	40822077	PLEKHH3	4.87
17	42392842	42392876	SLC25A39	8.34
17	79316557	79316689	TMEM105	6.57
18	5891215	5891256	TMEM200C	7.58
18	77398175	77398220	CTDP1	-4.22
19	3403888	3403934	NFIC	3.66
19	17219041	17219074	MYO9B	6.10
19	30744576	30744604	ZNF536	-7.87
19	34310911	34310951	KCTD15	-2.22
19	39798143	39798241	LRFN1	13.41
19	41107461	41107507	LTBP4	-3.87
19	44038525	44038544	ZNF575	12.28
19	45721137	45721153	EXOC3L2	-4.66
19	48698203	48698237	C19orf68	-14.12
19	48983836	48983869	LMTK3	6.33
19	51831025	51831166	VSIG10L	-5.66

19	55685105	55685146	SYT5	11.14
19	55944580	55944625	SHISA7	13.17
20	31045353	31045394	NOL4L	8.26
20	44803094	44803290	CDH22	2.84
20	56247612	56247651	PMEPA1	5.38
20	62463511	62463635	ZBTB46	3.01
22	17517094	17517125	CECR7	-11.77
22	19711581	19711627	GP1BB	4.88
22	38379681	38379718	SOX10	-16.43
22	38485555	38485598	BAIAP2L2	3.08
22	46500195	46500234	MIRLET7BHG	4.13
22	50355761	50355805	MIR6821	-1.37
22	51016280	51016323	CHKB	8.34

Supplemental Table 4. Demographic information for replication cohort

	Controls	Cases
N	18	17
Ethnicity*	9 (B); 9 (W)	10 (B); 7 (W)
Age (years)	37.6 ± 12.7	37.7 ± 6.8
PMI (hours)	20.4 ± 6.4	20.1 ± 6.4
pH	6.2 ± 0.4	6.4 ± 0.3
Sex**	18 (M)	17 (M)

*B= Black; W= White

** M=male

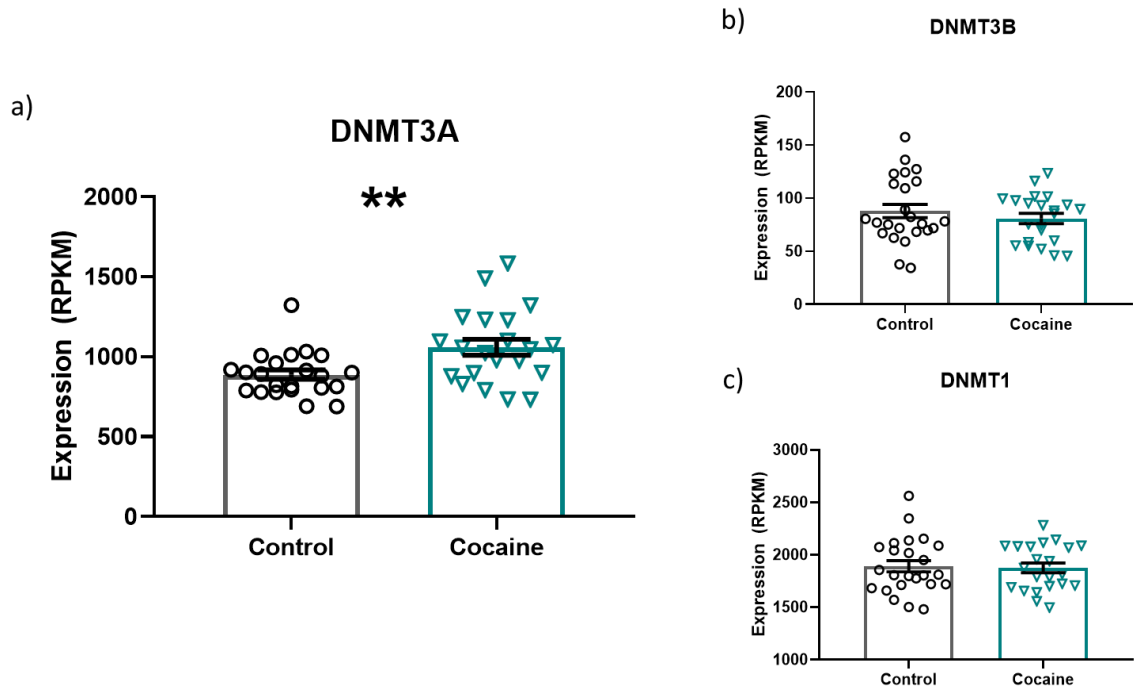
Supplemental Table 5. Targeted Bisulfite Amplicon Primers

Primer	Direction	Primer Sequence	Predicted Length
hg19_chr11:2188129-2188165			
TH_1	F	5' AAG TTY GTG YGT TTT GTA AGG 3'	269bp
	R	5' TCT ACA CCA CRC TAA AAA ACC TC 3'	
TH_2	F	5' GTT GTT YGT AGG AAG GAG GT 3'	216bp
	R	5' AAA CRC TTA ACT AAC CAT CCC 3'	

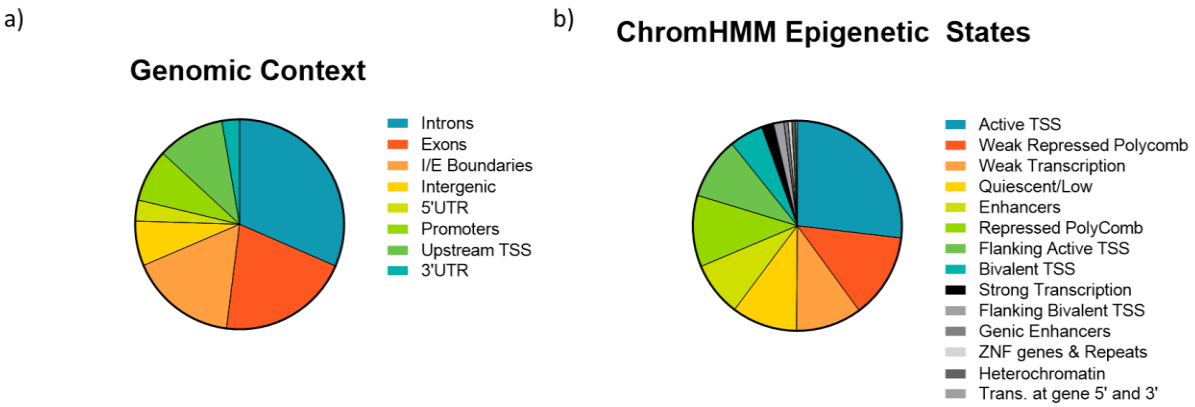
Supplemental Table 6. Luciferase Assay Primers

Primer	Direction	Primer Sequence
Human (hg19_chr11:2188129-2188165)		
a) THup	F	5' AGG CAT TAG AGG GCC CTG AGC CTG G 3'
b) THdw	R	5' ATA TAC TGG GTG CAC TGG AAC ACG C 3'
c) THAvrIIup	F	5' TAA TCC TAG GGG CAT TAG AGG GCC CTG AGC CTG G 3'
d) THSpe1dw	R	5' ATT AAC TAG TAT ATA CTG GGT GCA CTG GAA CAC GC 3'
e) pCPGfree-promMCSFw	F	5' CAC ACA CAT GTG TGC ATT CAT AAA TAT ATA C 3'
f) pCPGfree-promMCSrev	R	5' TTC TCA GGG ACT GTG GGC CAT GT 3'

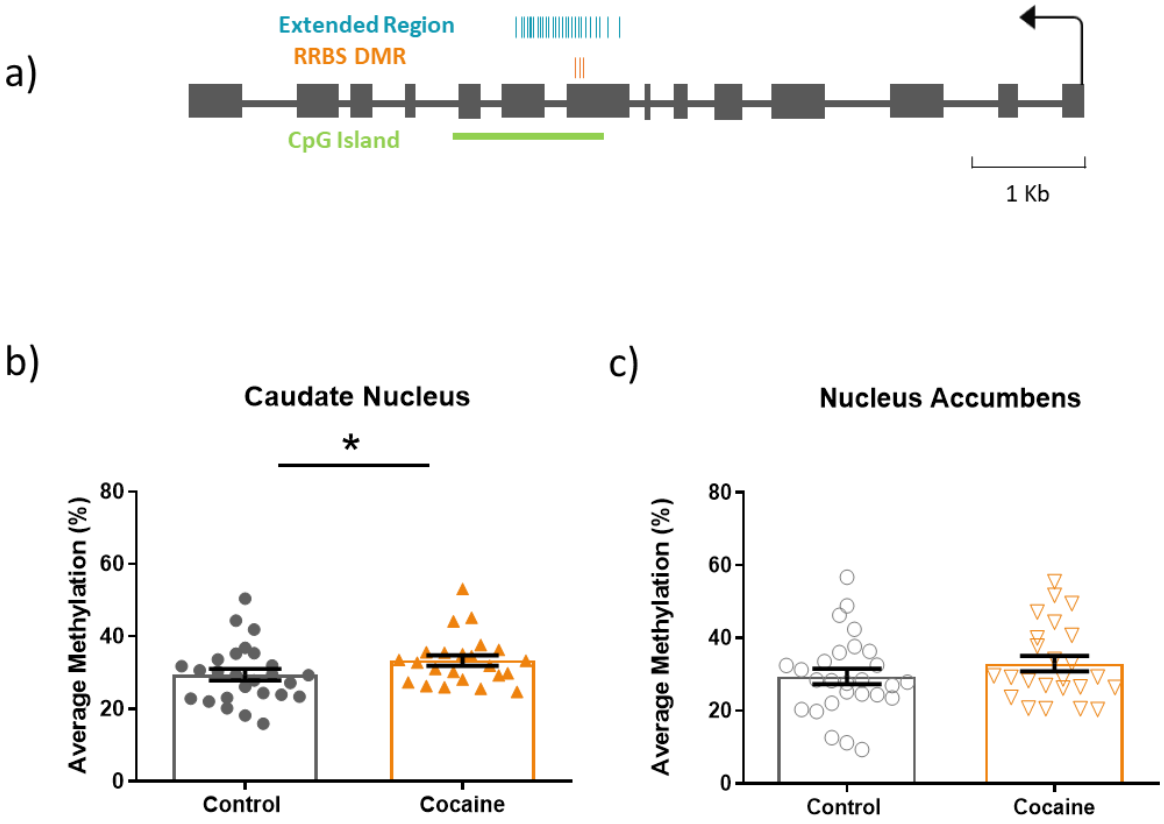
Figures



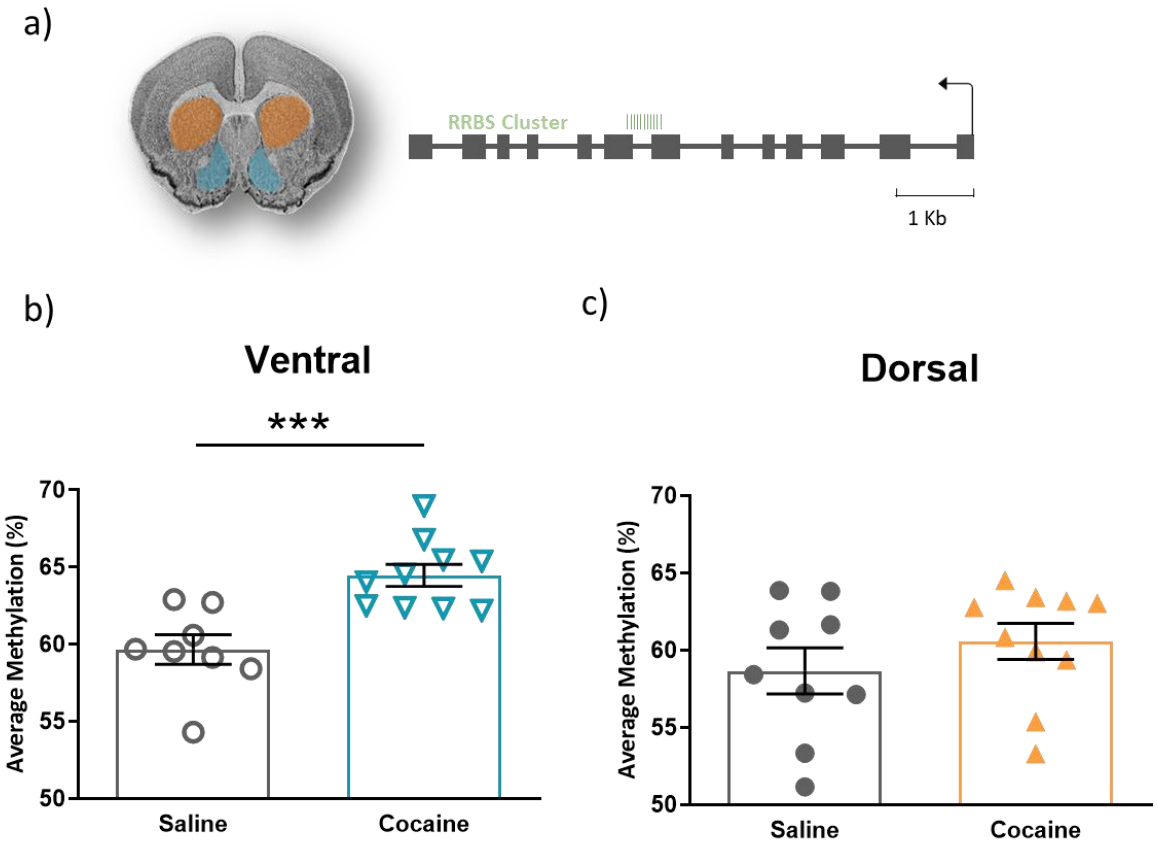
Supplementary Figure 1. *DNMT3A* gene expression is higher in the nucleus accumbens of cocaine dependent cases. a) Expression of the de novo methyltranferase *DNMT3A* is higher in cases (n=21) compared to controls (n=22). b) There is no difference between groups in terms of *DNMT3B* expression (n=22 cases and 24 controls) or c) *DNMT1* expression (n=23 cases and 24 controls) RPKM = reads per kilobase per million, Data represented as mean ± s.e.m., ** $p < 0.01$



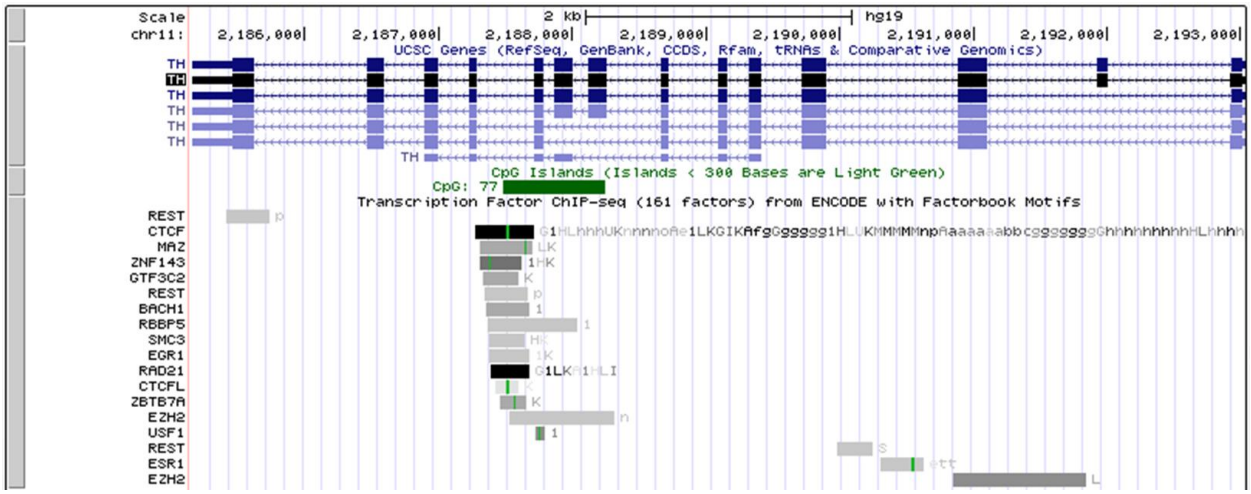
Supplementary Figure 2. Annotation of all CpG clusters used for differential methylation analysis. a) All 4814 CpG clusters distributed according to known genomic context and b) ChromHMM predicted epigenetic state.



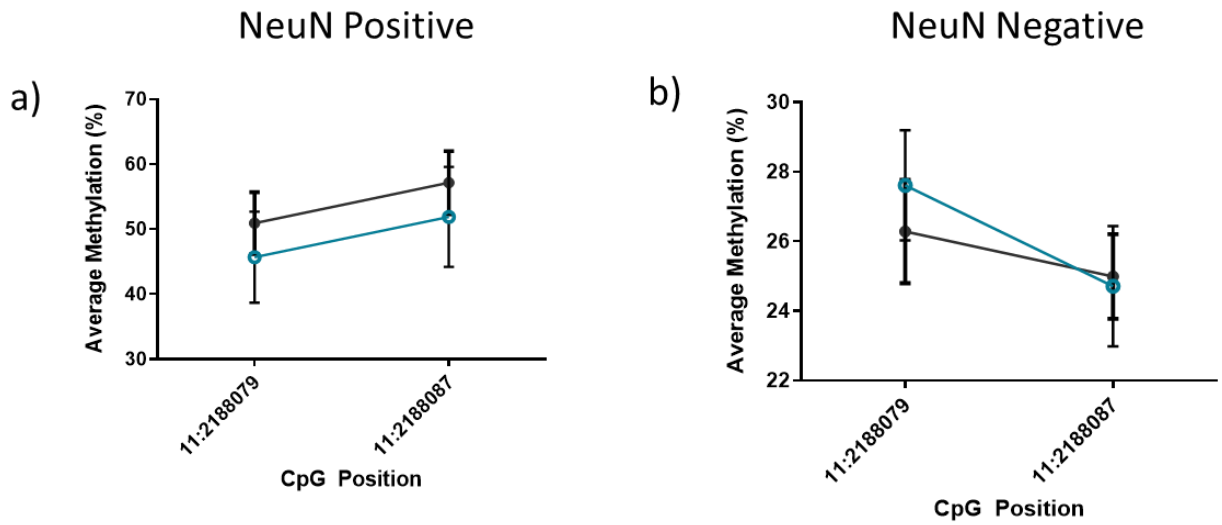
Supplementary Figure 3. Targeted bisulfite amplicon validation of the *TH* CpG cluster. a) Average methylation of the 3 CpGs within exon 8 of *TH* is higher in the caudate nucleus of cocaine dependent cases (n=23) compared to controls (n=26). b) We found no significant increase in the nucleus accumbens (n= 24 cases and 27 controls). Data represented as mean ± s.e.m., * $p < 0.05$



Supplementary Figure 4. Gene body methylation of *Th* in the striatum of cocaine self-administering mice. a) Genomic DNA from dorsal (orange) and ventral (blue) striatum was used to generate methylome data in mice, including a cluster of 11 CpGs with exon 8/9 of *Th*. b) This CpG cluster was hypermethylated in the ventral striatum (nucleus accumbens) of cocaine self-administering mice (n=10) compared to saline administered controls (n=8). c) No significant increase in average methylation was detected in the dorsal striatum (caudate-putamen; n= 10 cocaine animals and 9 controls)Data represented as mean ± s.e.m., *** $p < 0.001$



Supplementary Figure 5. UCSC Genome Browser view of known EGR1 ChIP-seq peak overlapping the intragenic CpG island of *TH*.



Supplementary Figure 6. No significant difference in EGR1 binding site methylation in FANS nuclei from the nucleus accumbens. a) We found no difference in NeuN+ nuclei, nor b) NeuN- nuclei (n=19 cases and 23 controls).

CHAPTER IV: DISCUSSION AND CONCLUDING REMARKS

1. Preface

The purpose of my doctoral thesis was to investigate the relationship between DNA methylation and cocaine dependence by examining the methylome of two key subregions of the human striatum. I began by generating the first methylome-wide datasets from the nucleus accumbens and caudate nucleus in samples from cocaine-dependent subjects, which provided two distinct, yet complementary avenues for deeper investigation. There is over 30 years of research that has linked chronic cocaine exposure, and its associated behavioral phenotypes, to lasting changes at the cellular level, and although epigenetic mechanisms have been thoroughly investigated in animal models, there is very little evidence that these phenomena are implicated in the human disorder. My results suggest that cocaine dependence is indeed associated with widespread alterations in DNA methylation in the human striatum and may regulate the expression of important genes through novel regulatory mechanisms. In this chapter, I will briefly review and integrate my findings into a theory for the role of DNA methylation in the pathophysiology of cocaine use disorder.

2. Discussion

2.1 Cocaine dependence is associated with widespread alterations of the striatal methylome

During the course of my doctoral degree, we used reduced representation bisulfite sequencing to generate the first human cocaine methylome datasets in two addiction-relevant subregions of the striatum: the nucleus accumbens and the caudate nucleus. This method allowed us to benefit from the relatively unbiased nature of next generation sequencing technologies, while still enriching our dataset for CpG-rich regions which are most likely to have regulatory potential (Deaton & Bird, 2011). In both brain areas, we identified over 100 clusters of CpGs which were differentially methylated between cases and controls. While we found distinct loci that were either more or less methylated in the cocaine group (discussed in chapters II and III, and below), we found that there was an over-representation of

hypermethylated loci in both brain areas. This finding was supported by our finding that gene expression levels of the de novo methyltransferase*, DNMT3A*, was heightened in both regions in the cocaine group as well. These results are well aligned with evidence from animals that finds differential expression of methyltransferases under disparate cocaine exposure paradigms, but sustained increases in *DNMT3a* after chronic drug seeking. We extend these finding to propose that the addition of methylation serves to solidify and maintain widespread transcriptomic changes that are induced by cocaine-related changes in cell signalling (C. Nagy, Vaillancourt, & Turecki, 2018). For example, while transient changes in histone modifications accompany acute exposures, methylation appears to be necessary for the maintenance of drug seeking behaviors.

We also found that intragenic regions appeared to be over-represented within our list of differentially methylated regions (DMRs) in both brain regions, with enhancers and transcription start sites being enriched in the caudate and accumbens datasets, respectively. Although RRBS enriches for CpG islands, which are canonically associated with promoters, our results highlight the functional importance of intragenic regions to brain health. While methylation at promoters is typically thought to decrease gene expression, intragenic methylation may serve more diverse functions including promoting alternative splicing events and gene expression through its oxidative products. Importantly, these are processes that have been recently implicated in cocaine-related behaviors in animals (Cates et al., 2018; Feng et al., 2014; Penrod et al., 2018). We also used gene ontology analysis to find that DMRs from both brain areas were closely associated with genes involved in DNA binding and transcriptional regulation. Given the importance of transcription factors in orchestrating cocaine-related transcriptional and behavioral changes in animals, this finding suggests that similar mechanisms are indeed implicated in the human disorder. It is likely that de novo DNA methylation events occur at these upstream regulatory factors in order to “prepare” the genome to respond to re-exposure events by priming or desensitizing critical genes.

2.2 IRX2 contains a novel regulatory element that is sensitive to cocaine in caudate neurons

When we sorted our DMR list in the caudate nucleus by size, we found a large cluster of CpGs within the third exon if *IRX2* to be less methylated in the cocaine group compared to controls. This gene codes for the Iroquois Homeobox 2 protein which is one of three transcriptional repressors in the *IRXA* gene family that are highly expressed in the developing nervous system and then become lowly expressed within the adult brain (Miller et al., 2017, Appendix 1. Figure 1). *IRX2* and its homolog, *IRX1* are in an evolutionarily conserved, head-to-head orientation, and although they share intergenic enhancers in animal models, their regulation was previously unstudied in human cells (Tena et al., 2011). Using chromatin conformation assays, we found that the two genes are in close physical contact in the human genome, especially in cells where the transcripts are expressed. Furthermore, we found that targeted methylation of the locus within exon 3 reduces the expression of both *IRX2* and *IRX1*, through attenuated CTCF binding and decreased chromatin contact.

Importantly, we found that the hypomethylation of *IRX2* in the caudate nucleus as associated with increased expression of the *IRX2* transcript in the cocaine group. While the downstream targets of this protein have yet to be experimentally validated, we used in silico prediction algorithms to identify potential target genes including two that has nominally significant gene expression changes in our dataset (Appendix 1. Figure 2). Interestingly, we found decreased expression of the *ADAM10* metalloprotease, which correlated with *IRX2* expression in the cocaine group alone (Appendix 1. Figure 3). Metalloprotease genes have previously been shown to be dysregulated in tissue from humans with drug use disorders, and ADAM10 in particular functions as an extracellular matrix remodelling complex in multiple brain disorders (Mash et al., 2007; Saftig & Lichtenthaler, 2015).

We posit that the cocaine dependence-associated activation of *IRX2* gene expression represents another example of drug use disorders reactivating neurodevelopmental transcriptional programs in

order to facilitate heightened states of plasticity. Developmental genes have been widely associated with maladaptive neuroplasticity in a number of brain disorders, but they may play an especially important role in the cellular and behavioral responses to chronic cocaine (Flores et al., 2005; Jassen, Yang, Miller, Calder, & Madras, 2006; Maheu & Ressler, 2017). For example, persistent activation of dopaminergic signalling increases the expression of deleted in colorectal cancer (*DCC*), an axon guidance molecule that is primarily expressed during neurodevelopment (Jassen et al., 2006). Furthermore, DCC knockout mice show attenuated behavioral responses to psychostimulants (Flores et al., 2005), which suggests that typically neurodevelopmental molecular cascades are necessary for drug dependent phenotypes. Although genetic manipulations of *IRX2* were beyond the scope of this thesis, we hypothesize that future studies that make use of in vivo epigenomic editors will find that altering methylation levels in caudate neurons will result in detectable perturbations in drug-related behaviors.

2.3 Cocaine-related intragenic methylation of tyrosine hydroxylase in the striatum

Given the remarkable similarity between the overall methylomic profiles of caudate nucleus and the nucleus accumbens, we next explored whether the two brain regions contained overlapping clusters of differential methylation from our RRBS studies. Interestingly, a cluster of 3 CpCs within the *TH* gene were hypermethylated in both brain areas in the cocaine group, and with a magnitude that is rarely observed in post-mortem studies. Since TH has a known relationship between reward-driven behaviors in animals (Kaminer et al., 2019; Logan et al., 2019), we used FANS separated nuclei, homogenate tissue from an independent cohort of samples, in vitro manipulations and supporting work from an animal model to explore this phenomenon further. We searched the region surrounding the cocaine DMR for known protein binding sites, and found a canonical EGR1 motif, a transcription factor that was previously shown to be induced by cocaine exposure and whose mRNA was elevated in both brain regions within the cocaine group of the current study. The binding site contains 2 CpGs which were hypermethylated in the caudate nucleus of the discovery cocaine group and the nucleus accumbens of

the replication cocaine group. Furthermore, we found the caudate hypermethylation to be specific to neuronal nuclei in our FANS experiment, and a similar increase in methylation within the nucleus accumbens of mice who were trained to self administer cocaine. Finally, we used an enhancer luciferase assay to show that the region containing the cocaine-related hypermethylation has the ability to act as an enhancer and that action is attenuated by methylation and enhanced by EGR1 expression.

Importantly, we were unable to detect any differences in *TH* expression within either brain region in this study which raises important considerations about our findings. As an enzymatic component of dopamine synthesis, *TH* is most abundantly expressed within dopaminergic neurons, and is often used as a marker of dopaminergic cell identity in immunohistochemical studies. Therefore, in the context of striatal tissue homogenates, a large proportion of *TH* RNA molecules are likely to be located within the afferent projections from midbrain cell bodies rather than cell bodies located within the striatum itself (Gervasi et al., 2016). Nevertheless, the epigenetic changes observed here suggest that chronic cocaine dependence has a functional impact on striatal *TH* gene regulation. In this context, our findings support two possible hypotheses: 1) midbrain TH expression at axon terminals remains unchanged after chronic cocaine and masks small changes in local striatal gene expression or 2) cocaine-induced overexpression of EGR1 and hypermethylation of the *TH* gene body work in opposition to keep gene expression within a homeostatic range.

In either case, the precise role of *TH* within striatal cells is unclear in healthy, let alone disease contexts. Recent work has identified a small subgroup of GABAergic interneurons within the striatum that express TH, are electrochemically distinct from other interneurons subgroups, and are involved in reward-driven behaviors but do not appear to be directly involved in dopamine synthesis (Ibáñez-Sandoval et al., 2010; Kaminer et al., 2019; Xenias et al., 2015). Although it is difficult to determine the relative abundance of these cells, a recent single cell transcriptome study detected *Th* transcripts in 4 out of 16 striatal cell types (Savell et al., 2020). Although *Th* expression remained unchanged in these

cell types after cocaine exposure, it is important to note that this study explored the effects of a single cocaine injection rather than an extended self-administration that would better mirror the chronic dependence displayed by human patients.

Given the heterogeneity of our findings and the reported rarity of *TH* expressing cells, we suggest that the novel regulatory element that we have identified here might have a strong impact on cocaine neurobiology, in a small number of cells. Although the availability of single-cell technologies is increasing rapidly, their functionality in post-mortem brain is continuing to prove challenging. Extended storage times and long post-mortem intervals are particularly damaging to the molecular integrity of these tissues, and its still unclear whether some cell types are more affected than others. Nonetheless, significant strides have been made in surveying the nuclear transcriptome of post-mortem cortical regions in humans, and single cell methylomic studies are certainly on the horizon (Corina Nagy et al., 2020). We are confident that these technologies will add to the identification and characterization of the most heavily impacted cell types in human cocaine dependence and will help clarify the roles of *TH* and *IRX2* in cocaine neurobiology.

2.4 Conclusions and Future Directions

Cocaine use disorder affects up to 1% of the North American population, and despite its considerable burden, our understanding of the pathophysiology of the disease in humans remains incomplete. Research in animals — which model distinct phases of drug exposure, drug seeking and withdrawal — have heavily implicated the mesocorticolimbic dopamine circuitry in cocaine-induced changes and neuroimaging studies in humans have largely corroborated these findings. Molecular research in these brain areas has revealed a role for epigenetic mechanisms, including DNA methylation, in translating long-term behavioral changes to lasting disruptions in cell functioning but little evidence has been presented in humans. This thesis contains the first methylome-wide investigations in humans,

with a focus on dependence-related changes in two downstream projection targets of midbrain dopamine: the nucleus accumbens and the caudate nucleus. Our findings support the hypothesis that cocaine dependence results in widespread alterations in striatal DNA methylation in humans, including at previously unexplored intragenic regulatory sites that have the potential to regulated addiction-related perturbations in cellular functioning. While the models proposed in this thesis are important steps towards fully understanding the interplay between cocaine dependence and epigenomics, important questions still remain about the specificity of these findings with regards to genomic context, cell type, and other important covariates.

For this work we focused on CpG methylation in males, as it represented the most well-understood mechanism of DNA methylation in the most readily available demographic, although we now know that methylation in non-CpG contexts is particularly important in brain development and health (Lister et al., 2013). Furthermore, cocaine dependence presents differently in males and females, even at the transcriptomic level (Savell et al., 2020), and future studies should include a range of biological sexes and genders to fully understand this disorder. Given the relatively small changes in methylation that we observed at each DMR, it is likely that large changes in rare cell populations have been masked by the tissue homogenates used in our studies — the use of single cell technologies in future studies is therefore paramount. In particular, future work should investigate transcriptomics and epigenomic changes within the same nuclei in order to fully understand the complex interactions between gene expression and multiple layers of the epigenome in all functional cell types.

Overall, we believe that DNA methylation has an important role in the etiology and outcome of cocaine use disorder in humans, including at the *IRX2* and *TH* genes. Through identifying altered methylation signatures, we are better positioned to understand the neurobiology of cocaine use disorder and although additional investigations are necessary, modifying methylation dynamics at key transcriptomic regulators presents a promising opportunity for pharmaceutical intervention.

References

Adlaf, E. M., Begin, P., & Sawka, E. (2005). *Canadian Addiction Survey (CAS)*.

Aharonovich, E., Hasin, D. S., Brooks, A. C., Liu, X., Bisaga, A., & Nunes, E. V. (2006). Cognitive deficits predict low treatment retention in cocaine dependent patients. *Drug and Alcohol Dependence*, *81*(3), 313–322. https://doi.org/10.1016/j.drugalcdep.2005.08.003

Alaghband, Y., Bredy, T. W., & Wood, M. A. (2016). The role of active DNA demethylation and Tet enzyme function in memory formation and cocaine action. *Neuroscience Letters*, *625*, 40–46. https://doi.org/10.1016/j.neulet.2016.01.023

Albein-Urios, N., Verdejo-Román, J., Asensio, S., Soriano-Mas, C., Martínez-González, J. M., & Verdejo-García, A. (2014). Re-appraisal of negative emotions in cocaine dependence: Dysfunctional corticolimbic activation and connectivity. *Addiction Biology*, *19*(3), 415–426. https://doi.org/10.1111/j.1369-1600.2012.00497.x

Albertson, D. N., Pruetz, B., Schmidt, C. J., Kuhn, D. M., Kapatos, G., & Bannon, M. J. (2004). Gene expression profile of the nucleus accumbens of human cocaine abusers: evidence for dysregulation of myelin. *Journal of Neurochemistry*, *88*. https://doi.org/10.1046/j.1471-4159.2003.02247.x

Albertson, D., Pruetz, B., Schmidt, C., Kuhn, D., Kapatos, G., & Bannon, M. (2004). Gene expression profile of the nucleus accumbens of human cocaine abusers: evidence for dysregulation of myelin. *Journal of Neurochemistry*, *88*(5), 1211–1219. https://doi.org/10.1046/j.1471-4159.2003.02247.x

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Anier, K., Malinovskaja, K., Aonurm-Helm, A., Zharkovsky, A., Kalda, A., & Anu, A.-H. (2010). DNA methylation regulates cocaine-induced behavioral sensitization in mice. *Neuropsychopharmacology : Official Publication of the American College of 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. *The International Journal of Neuropsychopharmacology / Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, *16*(9), 2053–2066. https://doi.org/10.1017/S1461145713000394

Ausió, J. (2016). MeCP2 and the enigmatic organization of brain chromatin. Implications for depression and cocaine addiction. *Clinical Epigenetics*, *8*, 58. https://doi.org/10.1186/s13148-016-0214-5

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., & Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nature Chemistry*, *6*(12), 1049–1055. https://doi.org/10.1038/nchem.2064

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

Bannon, M. J., Johnson, M. M., Michelhaugh, S. K., Hartley, Z. J., Halter, S. D., David, J. A., … 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

Bannon, M. J., Savonen, C. L., Jia, H., Dachet, F., Halter, S. D., Schmidt, C. J., … Kapatos, G. (2015). Identification of long noncoding RNAs dysregulated in the midbrain of human cocaine abusers. *Journal of Neurochemistry*, *135*, 50–59. https://doi.org/10.1111/jnc.13255

Barak Caine, S., Heinrichs, S. C., Coffin, V. L., & Koob, G. F. (1995). Effects of the dopamine D-1 antagonist SCH 23390 microinjected into the accumbens, amygdala or striatum on cocaine self-administration in the rat. *Brain Research*, *692*(1–2), 47–56. https://doi.org/10.1016/0006-8993(95)00598-K

Barros, M., Dempster, E. L., Illott, N., Chabrawi, S., Maior, R. S., Tomaz, C., … Müller, C. P. (2013). Decreased methylation of the NK3 receptor coding gene (TACR3) after cocaine-induced place preference in marmoset monkeys. *Addiction Biology*, *18*(3), 452–454. https://doi.org/10.1111/j.1369-1600.2011.00409.x

Belin, D., & Everitt, B. (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/10.1016/j.neuron.2007.12.019

Beveridge, T. J. R., Smith, H. R., Nader, M. a, & Porrino, L. J. (2005). Effects of chronic cocaine self-administration on norepinephrine transporters in the nonhuman primate brain. *Psychopharmacology*, *180*(4), 781–788. https://doi.org/10.1007/s00213-005-2162-1

Bird, A. P., & Wolffe, A. P. (1999). Methylation-Induced Repression— Belts, Braces, and Chromatin. *Cell*, *99*(5), 451–454. https://doi.org/10.1016/S0092-8674(00)81532-9

Bucello, C., Degenhardt, L., Calabria, B., Nelson, P., Roberts, A., Medina-Mora, M., & Compton, W. (2010). What do we know about the extent of cocaine use and dependence? Results of a global systematic review. *NDARC Technical Report No. 308*. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0

Caine, S. B., Thomsen, M., Gabriel, K. I., Berkowitz, J. S., Gold, L. H., Koob, G. F., … Xu, M. (2007). Lack of self-administration of cocaine in dopamine D1 receptor knock-out mice. *Journal of Neuroscience*, *27*(48), 13140–13150. https://doi.org/10.1523/JNEUROSCI.2284-07.2007

Calu, D. J., Stalnaker, T. A., Franz, T. M., Singh, T., Shaham, Y., & Schoenbaum, G. (2007). Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. *Learning and Memory*, *14*(5), 325–328. https://doi.org/10.1101/lm.534807

Caputi, F. F., Di Benedetto, M., Carretta, D., Bastias del Carmen Candia, S., D'Addario, C., Cavina, C., … Romualdi, P. (2014). Dynorphin/KOP and nociceptin/NOP gene expression and epigenetic changes by cocaine in rat striatum and nucleus accumbens. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *49*, 36–46. https://doi.org/10.1016/j.pnpbp.2013.10.016

Carlezon, W. A., Thome, J., Olson, V. G., Lane-Ladd, S. B., Brodkin, E. S., Hiroi, N., … Nestler, E. J. (1998). Regulation of cocaine reward by CREB. *Science*, *282*(5397), 2272–2275. https://doi.org/10.1126/science.282.5397.2272

Carouge, D., Host, L., Aunis, D., Zwiller, J., & Anglard, P. (2010). CDKL5 is a brain MeCP2 target gene regulated by DNA methylation. *Neurobiology of Disease*, *38*(3), 414–424.

https://doi.org/10.1016/j.nbd.2010.02.014

Cassel, S., Carouge, D., Gensburger, C., Anglard, P., Burgun, C., Dietrich, J.-B., … Zwiller, J. (2006). Fluoxetine and cocaine induce the epigenetic factors MeCP2 and MBD1 in adult rat brain. *Molecular Pharmacology*, *70*(2), 487–492. https://doi.org/10.1124/mol.106.022301

Cates, H. M., Heller, E. A., Lardner, C. K., Purushothaman, I., Peña, C. J., Walker, D. M., … Nestler, E. J. (2018). Transcription Factor E2F3a in Nucleus Accumbens Affects Cocaine Action via Transcription and Alternative Splicing. *Biological Psychiatry*, *84*(3), 167–179. https://doi.org/10.1016/j.biopsych.2017.11.027

Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics*, *10*(5), 295–304. https://doi.org/10.1038/nrg2540

Chandra, R., Francis, T. C., Konkalmatt, P., Amgalan, A., Gancarz, A. M., Dietz, D. M., & Lobo, M. K. (2015). Opposing Role for Egr3 in Nucleus Accumbens Cell Subtypes in Cocaine Action. *Journal of Neuroscience*, *35*(20), 7927–7937. https://doi.org/10.1523/JNEUROSCI.0548-15.2015

Charlson, F. J., Baxter, A. J., Dua, T., Degenhardt, L., Whiteford, H. A., & Vos, T. (2015). Excess mortality from mental, neurological and substance use disorders in the Global Burden of Disease Study 2010. *Epidemiology and Psychiatric Sciences*, *24*(2), 121–140. https://doi.org/10.1017/S2045796014000687

Contreras-Rodríguez, O., Albein-Urios, N., Vilar-López, R., Perales, J. C., Martínez-Gonzalez, J. M., Fernández-Serrano, M. J., … Verdejo-García, A. (2016). Increased corticolimbic connectivity in cocaine dependence versus pathological gambling is associated with drug severity and emotion-related impulsivity. *Addiction Biology*, *21*(3), 709–718. https://doi.org/10.1111/adb.12242

Cowan, C. W., Taniguchi, M., & Hale, C. F. (2008). Regulation of synaptic connectivity with chronic cocaine. *The American Journal of Psychiatry*, *165*(11), 1393. https://doi.org/10.1176/appi.ajp.2008.08101465

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes and Development*, *25*(10), 1010–1022. https://doi.org/10.1101/gad.2037511

Degenhardt, L., & Hall, W. (2012, January 7). Extent of illicit drug use and dependence, and their contribution to the global burden of disease. *The Lancet*. Elsevier. https://doi.org/10.1016/S0140-6736(11)61138-0

Delfs, J. M., Zhu, Y., Druhan, J. P., & Aston-Jones, G. (2000). Noradrenaline in the ventral forebrain is critical for opiate withdrawal-induced aversion. *Nature*, *403*(6768), 430–434. https://doi.org/10.1038/35000212

Deng, J. V, Rodriguiz, R. M., Hutchinson, A. N., Kim, I.-H., Wetsel, W. C., & West, A. E. (2010). MeCP2 in the nucleus accumbens contributes to neural and behavioral responses to psychostimulants. *Nature Neuroscience*, *13*(9), 1128–1136. https://doi.org/10.1038/nn.2614

Dong, Y. (2016). Silent synapse-based circuitry remodeling in drug addiction. *International Journal of Neuropsychopharmacology*. Oxford University Press. https://doi.org/10.1093/ijnp/pyv136

Dumitriu, D., Laplant, Q., Grossman, Y. S., Dias, C., Janssen, W. G., Russo, S. J., … Nestler, E. J. (2012). Subregional, dendritic compartment, and spine subtype specificity in cocaine regulation of dendritic spines in the nucleus accumbens. *The Journal of Neuroscience : The Official Journal of the*

Society for Neuroscience, *32*(20), 6957–6966. https://doi.org/10.1523/JNEUROSCI.5718-11.2012

Enoch, M.-A., Baghal, B., Yuan, Q., & Goldman, D. (2013). A factor analysis of global GABAergic gene expression in human brain identifies specificity in response to chronic alcohol and cocaine exposure. *PloS One*, *8*(5), e64014. https://doi.org/10.1371/journal.pone.0064014

Enoch, M., Rosser, a a, Zhou, Z., Mash, D. C., Yuan, Q., & Goldman, D. (2014). Expression of glutamatergic genes in healthy humans across 16 brain regions; altered expression in the hippocampus after chronic exposure to alcohol or cocaine. *Genes, Brain, and Behavior*, *13*(8), 758–768. https://doi.org/10.1111/gbb.12179

Everitt, B. J., & Robbins, T. W. (2005, November). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*. Nat Neurosci. https://doi.org/10.1038/nn1579

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

Farrelly, L. A., Thompson, R. E., Zhao, S., Lepack, A. E., Lyu, Y., Bhanu, N. V., … Maze, I. (2019). Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3. *Nature*, *567*(7749), 535–539. https://doi.org/10.1038/s41586-019-1024-7

Fatemi, M., Hermann, A., Gowher, H., & Jeltsch, A. (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *European Journal of Biochemistry*, *269*(20), 4981–4984. https://doi.org/10.1046/j.1432-1033.2002.03198.x

Feng, J., Shao, N., Szulwach, K. E., Vialou, V., Huynh, J., Zhong, C., … Nestler, E. J. (2015). Role of Tet1 and 5-hydroxymethylcytosine in cocaine action. *Nature Neuroscience*, *18*(4), 536–544. https://doi.org/10.1038/nn.3976

Feng, J., Wilkinson, M., Liu, X., Purushothaman, I., Ferguson, D., Vialou, V., … Shen, L. (2014). Chronic cocaine-regulated epigenomic changes in mouse nucleus accumbens. *Genome Biology*, *15*(4), R65. https://doi.org/10.1186/gb-2014-15-4-r65

Flores, C., Manitt, C., Rodaros, D., Thompson, K. M., Rajabi, H., Luk, K. C., … Kennedy, T. E. (2005). Netrin receptor deficient mice exhibit functional reorganization of dopaminergic systems and do not sensitize to amphetamine. *Molecular Psychiatry*, *10*(6), 606–612. https://doi.org/10.1038/sj.mp.4001607

Fonteneau, M., Filliol, D., Anglard, P., Befort, K., Romieu, P., & Zwiller, J. (2016). Inhibition of DNA methyltransferases regulates cocaine self-administration by rats: a genome-wide DNA methylation study. *Genes, Brain and Behavior*, 1–15. https://doi.org/10.1111/gbb.12354

Fragou, D., Zanos, P., Kouidou, S., Njau, S., Kitchen, I., Bailey, A., & Kovatsi, L. (2013). Effect of chronic heroin and cocaine administration on global DNA methylation in brain and liver. *Toxicology Letters*, *218*(3), 260–265. https://doi.org/10.1016/j.toxlet.2013.01.022

Freeman, W. M., Lull, M. E., Patel, K. M., Brucklacher, R. M., Morgan, D., Roberts, D. C. S., & Vrana, K. E. (2010). Gene expression changes in the medial prefrontal cortex and nucleus accumbens following abstinence from cocaine self-administration. *BMC Neuroscience*, *11*, 29. https://doi.org/10.1186/1471-2202-11-29

Freeman, W., Nader, M., Nader, S., Robertson, D., Gioia, L., Mitchell, S., … Vrana, K. (2001). Chronic cocaine-mediated changes in non-human primate nucleus accumbens gene expression. *Journal of Neurochemistry*, *77*(2), 542–549. https://doi.org/10.1046/j.1471-4159.2001.00382.x

Gervasi, N. M., Scott, S. S., Aschrafi, A., Gale, J., Vohra, S. N., Macgibeny, M. A., … Kaplan, B. B. (2016). The local expression and trafficking of tyrosine hydroxylase mRNA in the axons of sympathetic neurons. *RNA*, *22*(6), 883–895. https://doi.org/10.1261/rna.053272.115

Gigek, C. O., Chen, E. S., & Smith, M. A. C. (2016). Methyl-CpG-Binding Protein (MBD) Family: Epigenomic Read-Outs Functions and Roles in Tumorigenesis and Psychiatric Diseases. *Journal of Cellular Biochemistry*, *117*(1), 29–38. https://doi.org/10.1002/jcb.25281

Global Burden of Disease Collaborative Network. (n.d.). Global Burden of Disease Study (2019) Results. Retrieved November 22, 2020, from http://ghdx.healthdata.org/gbd-results-tool

Goldstein, R. Z., & Volkow, N. D. (2011, November). Dysfunction of the prefrontal cortex in addiction: Neuroimaging findings and clinical implications. *Nature Reviews Neuroscience*. NIH Public Access. https://doi.org/10.1038/nrn3119

Gross, J. A., Pacis, A., Chen, G. G., Barreiro, L. B., Ernst, C., & Turecki, G. (2015). Characterizing 5-hydroxymethylcytosine in human prefrontal cortex at single base resolution. *BMC Genomics*, *16*(1), 672. https://doi.org/10.1186/s12864-015-1875-8

Guan, X., & Guan, Y. (2013). Proteomic profile of differentially expressed proteins in the medial prefrontal cortex after repeated cocaine exposure. *Neuroscience*, *236C*, 262–270. https://doi.org/10.1016/j.neuroscience.2013.01.039

Guo, J. U., Su, Y., Shin, J. J. H., Shin, J. J. H., Li, H., Xie, B., … Song, H. (2013). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature Neuroscience*, *17*(2), 215–222. https://doi.org/10.1038/nn.3607

Guo, J. U., Su, Y., Zhong, C., Ming, G., & Song, H. (2011). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle (Georgetown, Tex.)*, *10*(16), 2662–2668. https://doi.org/10.4161/cc.10.16.17093

Guo, M.-L., Periyasamy, P., Liao, K., Kook, Y. H., Niu, F., Callen, S. E., & Buch, S. (2016). Cocaine-mediated downregulation of microglial miR-124 expression involves promoter DNA methylation. *Epigenetics*, *11*(11), 819–830. https://doi.org/10.1080/15592294.2016.1232233

Han, J., Li, Y., Wang, D., Wei, C., Yang, X., & Sui, N. (2010). Effect of 5-aza-2-deoxycytidine microinjecting into hippocampus and prelimbic cortex on acquisition and retrieval of cocaine-induced place preference in C57BL/6 mice. *European Journal of Pharmacology*, *642*(1–3), 93–98. https://doi.org/10.1016/j.ejphar.2010.05.050

Hasin, D. S., O'Brien, C. P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., … Grant, B. F. (2013, August 1). DSM-5 criteria for substance use disorders: Recommendations and rationale. *American Journal of Psychiatry*. American Psychiatric Association. https://doi.org/10.1176/appi.ajp.2013.12060782

Heal, D. J., Gosden, J., & Smith, S. L. (2014). Dopamine reuptake transporter (DAT) "inverse agonism" – A novel hypothesis to explain the enigmatic pharmacology of cocaine. *Neuropharmacology*, *87*, 19–40. https://doi.org/10.1016/j.neuropharm.2014.06.012

Heiman, M., Schaefer, A., Gong, S., Peterson, J. D., Day, M., Ramsey, K. E., … Heintz, N. (2008). A translational profiling approach for the molecular characterization of CNS cell types. *Cell*, *135*(4), 738–748. https://doi.org/10.1016/j.cell.2008.10.028

Hiroi, N., Fienberg, A. A., Haile, C. N., Alburges, M., Hanson, G. R., Greengard, P., & Nestler, E. J. (1999). Neuronal and behavioural abnormalities in striatal function in DARPP-32-mutant mice. *The European Journal of Neuroscience*, *11*(3), 1114–1118.

Hope, B. T. (1998). Cocaine and the AP-1 transcription factor complex. *Annals of the New York Academy of Sciences*, *844*, 1–6.

Host, L., Dietrich, J.-B., Carouge, D., Aunis, D., & Zwiller, J. (2011). Cocaine self-administration alters the expression of chromatin-remodelling proteins; modulation by histone deacetylase inhibition. *Journal of Psychopharmacology {(Oxford,} England)*, *25*(2), 222–229. https://doi.org/10.1177/0269881109348173

Hsieh, C. L. (1999). In vivo activity of murine de novo methyltransferases, Dnmt3a and Dnmt3b. *Molecular and Cellular Biology*, *19*(12), 8211–8218. https://doi.org/10.1128/MCB.19.12.8211

Hurd, Y. L., & Herkenham, M. (1993). Molecular alterations in the neostriatum of human cocaine addicts. *Synapse (New York, N.Y.)*, *13*(4), 357–369. https://doi.org/10.1002/syn.890130408

Ibáñez-Sandoval, O., Tecuapetla, F., Unal, B., Shah, F., Koós, T., & Tepper, J. M. (2010). Electrophysiological and morphological characteristics and synaptic connectivity of tyrosine hydroxylase-expressing neurons in adult mouse striatum. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(20), 6999–7016. https://doi.org/10.1523/JNEUROSCI.5996-09.2010

Im, H.-I., Hollander, J. A., Bali, P., & Kenny, P. J. (2010). MeCP2 controls BDNF expression and cocaine intake through homeostatic interactions with microRNA-212. *Nature Neuroscience*, *13*(9), 1120–1127. https://doi.org/10.1038/nn.2615

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., … Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (New York, N.Y.)*, *333*(6047), 1300–1303. https://doi.org/10.1126/science.1210597

Jassen, A. K., Yang, H., Miller, G. M., Calder, E., & Madras, B. K. (2006). Receptor regulation of gene expression of axon guidance molecules: Implications for adaptation. *Molecular Pharmacology*, *70*(1), 71–77. https://doi.org/10.1124/mol.105.021998

Jeltsch, A., & Jurkowska, R. Z. (2014). New concepts in DNA methylation. *Trends in Biochemical Sciences*, *39*(7), 310–318. https://doi.org/10.1016/j.tibs.2014.05.002

Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. *Science (New York, N.Y.)*, *293*(5532), 1074–1080. https://doi.org/10.1126/science.1063127

Kaminer, J., Espinoza, D., Bhimani, S., Tepper, J. M., Koos, T., & Shiflett, M. W. (2019). Loss of striatal tyrosine-hydroxylase interneurons impairs instrumental goal-directed behavior. *European Journal of Neuroscience*, *50*(4), 2653–2662. https://doi.org/10.1111/ejn.14412

Kennedy, P. J., Feng, J., Robison, a J., Maze, I., Badimon, A., Mouzon, E., … Nestler, E. J. (2013). Class I HDAC Inhibition Blocks Cocaine-Induced Plasticity Through Targeted Changes in Histone Methylation. *Nature Neuroscience*, *16*(4), 434–440. https://doi.org/10.1038/nn.3354.Class

Khatri, P., & Drăghici, S. (2005). Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics (Oxford, England)*, *21*(18), 3587–3595. https://doi.org/10.1093/bioinformatics/bti565

Koob, G. (2013). Negative reinforcement in drug addiction: the darkness within. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2013.03.011

Koob, G. F. (2008, July 10). A Role for Brain Stress Systems in Addiction. *Neuron*. NIH Public Access. https://doi.org/10.1016/j.neuron.2008.06.012

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

Kozlenkov, A., Roussos, P., Timashpolsky, A., Barbu, M., Rudchenko, S., Bibikova, M., … Dracheva, S. (2014). Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. *Nucleic Acids Research*, *42*(1), 109–127. https://doi.org/10.1093/nar/gkt838

Kozlenkov, A., Wang, M., Roussos, P., Rudchenko, S., Barbu, M., Bibikova, M., … Dracheva, S. (2015). Substantial DNA methylation differences between two major neuronal subtypes in human brain. *Nucleic Acids Research*, *44*(6), 2593–2612. https://doi.org/10.1093/nar/gkv1304

Kristiansen, L. V., Bannon, M. J., & Meador-Woodruff, J. H. (2009). Expression of Transcripts for Myelin Related Genes in Postmortem Brain from Cocaine Abusers. *Neurochemical Research*, *34*(1), 46–54. https://doi.org/10.1007/s11064-008-9655-3

Kumar, A., Choi, K.-H., Renthal, W., Tsankova, N. M., Theobald, D. E. H., Truong, H.-T., … Nestler, E. J. (2005). Chromatin remodeling is a key mechanism underlying cocaine-induced plasticity in striatum. *Neuron*, *48*(2), 303–314. https://doi.org/10.1016/j.neuron.2005.09.023

Kundu, K., & Rao, M. R. . (1999). CpG Islands in Chromatin Organization and Gene Expression. *Journal of Biochemistry*, *125*(2), 217–222.

Lambert, B. L., & Bauer, C. R. (2012). Developmental and behavioral consequences of prenatal cocaine exposure: a review. *Journal of Perinatology : Official Journal of the California Perinatal Association*, *32*(11), 819–828. https://doi.org/10.1038/jp.2012.90

LaPlant, Q., Vialou, V., Covington, H. E., Dumitriu, D., Feng, J., Warren, B. L., … Nestler, E. J. (2010). Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. *Nature Neuroscience*, *13*(9), 1137–1143. https://doi.org/10.1038/nn.2619

Larson, E. B., Graham, D. L., Arzaga, R. R., Buzin, N., Webb, J., Green, T. A., … Self, D. W. (2011). Overexpression of CREB in the nucleus accumbens shell increases cocaine reinforcement in self-administering rats. *Journal of Neuroscience*, *31*(45), 16447–16457. https://doi.org/10.1523/JNEUROSCI.3070-11.2011

Lee, B. R., Ma, Y. Y., Huang, Y. H., Wang, X., Otaka, M., Ishikawa, M., … Dong, Y. (2013). Maturation of silent synapses in amygdala-accumbens projection contributes to incubation of cocaine craving. *Nature Neuroscience*, *16*(11), 1644–1651. https://doi.org/10.1038/nn.3533

Lepack, A. E., Werner, C. T., Stewart, A. F., Fulton, S. L., Zhong, P., Farrelly, L. A., … Maze, I. (2020). Dopaminylation of histone H3 in ventral tegmental area regulates cocaine seeking. *Science*, *368*(6487), 197–201. https://doi.org/10.1126/science.aaw8806

Lev Maor, G., Yearim, A., & Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. *Trends in Genetics*, *31*(5), 274–280. https://doi.org/10.1016/j.tig.2015.03.002

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet*, *3*(9), 662–673. https://doi.org/10.1038/nrg887

Lister, R., Mukamel, E. a, Nery, J. R., Urich, M., Puddifoot, C. a, Johnson, N. D., … Ecker, J. R. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science (New York, N.Y.)*, *341*(6146), 1237905. https://doi.org/10.1126/science.1237905

Logan, R. W., Parekh, P. K., Kaplan, G. N., Becker-Krail, D. D., Williams, W. P., Yamaguchi, S., … McClung, C. A. (2019). NAD+ cellular redox and SIRT1 regulate the diurnal rhythms of tyrosine hydroxylase and conditioned cocaine reward. *Molecular Psychiatry*, *24*(11), 1668–1684. https://doi.org/10.1038/s41380-018-0061-1

Ma, Y. Y., Lee, B. R., Wang, X., Guo, C., Liu, L., Cui, R., … Dong, Y. (2014). Bidirectional modulation of incubation of cocaine craving by silent synapse-based remodeling of prefrontal cortex to accumbens projections. *Neuron*, *83*(6), 1453–1467. https://doi.org/10.1016/j.neuron.2014.08.023

Maheu, M. E., & Ressler, K. J. (2017, September 1). Developmental pathway genes and neural plasticity underlying emotional learning and stress-related disorders. *Learning and Memory*. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/lm.044271.116

Malvaez, M., Susan, M., Rogge, G., Astarabadi, M., Jacques, V., Carreiro, S., … Wood, M. (2013). {HDAC3-selective} inhibitor enhances extinction of cocaine-seeking behavior in a persistent manner. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(7), 2647–2652. https://doi.org/10.1073/pnas.1213364110

Mao, L., Horton, E., Guo, M., Xue, B., Jin, D., Fibuch, E. E., & Wang, J. Q. (2012). Cocaine increases phosphorylation of MeCP2 in the rat striatum in vivo : a differential role of NMDA receptors, *59*(5), 610–617. https://doi.org/10.1016/j.neuint.2011.04.013.Cocaine

Mash, D. C., Ffrench-Mullen, J., Adi, N., Qin, Y., Buck, A., & Pablo, J. (2007). Gene Expression in Human Hippocampus from Cocaine Abusers Identifies Genes which Regulate Extracellular Matrix Remodeling. *PLoS ONE*, *2*(11). https://doi.org/10.1371/journal.pone.0001187

Mash, D., Ffrench-Mullen, J., Adi, N., Qin, Y., Buck, A., Pablo, J., & ffrench-Mullen, Jarlath. (2007). Gene expression in human hippocampus from cocaine abusers identifies genes which regulate extracellular matrix remodeling. *{PloS} One*, *2*(11). https://doi.org/10.1371/journal.pone.0001187

Massart, R., Barnea, R., Dikshtein, Y., Suderman, M., Meir, O., Hallett, M., … Yadid, G. (2015). Role of DNA Methylation in the Nucleus Accumbens in Incubation of Cocaine Craving. *Journal of Neuroscience*, *35*(21), 8042–8058. https://doi.org/10.1523/JNEUROSCI.3053-14.2015

Massart, Renaud, Barnea, R., Dikshtein, Y., Suderman, M., Meir, O., Hallett, M., … Yadid, G. (2015). Role of DNA methylation in the nucleus accumbens in incubation of cocaine craving. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *35*(21), 8042–8058. https://doi.org/10.1523/JNEUROSCI.3053-14.2015

Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., … Costello, J. F. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, *466*(7303), 253–257. https://doi.org/10.1038/nature09165

Maze, I., Covington, H. E., Dietz, D. M., LaPlant, Q., Renthal, W., Russo, S. J., … Nestler, E. J. (2010). Essential role of the histone methyltransferase G9a in cocaine-induced plasticity. *Science (New York, N.Y.)*, *327*(5962), 213–216. https://doi.org/10.1126/science.1179438

Maze, I., Feng, J., Wilkinson, M. B., Sun, H., Shen, L., & Nestler, E. J. (2011). Cocaine dynamically regulates heterochromatin and repetitive element unsilencing in nucleus accumbens. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3035–3040. https://doi.org/10.1073/pnas.1015483108

McBride, W. J., Murphy, J. M., & Ikemoto, S. (1999). Localization of brain reinforcement mechanisms: Intracranial self- administration and intracranial place-conditioning studies. *Behavioural Brain Research*, *101*(2), 129–152. https://doi.org/10.1016/S0166-4328(99)00022-4

McClernon, F. J., Kozink, R. V., Lutz, A. M., & Rose, J. E. (2009). 24-h smoking abstinence potentiates fMRI-BOLD activation to smoking cues in cerebral cortex and dorsal striatum. *Psychopharmacology*, *204*(1), 25–35. https://doi.org/10.1007/s00213-008-1436-9

McClung, C. A., & Nestler, E. J. (2003). Regulation of gene expression and cocaine reward by CREB and ΔFosB. *Nature Neuroscience*, *6*(11), 1208–1215. https://doi.org/10.1038/nn1143

McGregor, A., & Roberts, D. C. S. (1993). Dopaminergic antagonism within the nucleus accumbens or the amygdala produces differential effects on intravenous cocaine self-administration under fixed and progressive ratio schedules of reinforcement. *Brain Research*, *624*(1–2), 245–252. https://doi.org/10.1016/0006-8993(93)90084-Z

Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., & Heintz, N. (2012). MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. *Cell*, *151*(7), 1417–1430. https://doi.org/10.1016/j.cell.2012.11.022

Miller, J. A., Guillozet-Bongaarts, A., Gibbons, L. E., Postupna, N., Renz, A., Beller, A. E., … Lein, E. (2017). Neuropathological and transcriptomic characteristics of the aged brain. *ELife*, *6*. https://doi.org/10.7554/eLife.31126

Mo, A., Mukamel, E. A., Davis, F. P., Luo, C., Henry, G. L., Picard, S., … Nathans, J. (2015). Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron*, *86*(6), 1369–1384. https://doi.org/10.1016/j.neuron.2015.05.018

Muller Ewald, V. A., & LaLumiere, R. T. (2018, November 1). Neural systems mediating the inhibition of cocaine-seeking behaviors. *Pharmacology Biochemistry and Behavior*. Elsevier Inc. https://doi.org/10.1016/j.pbb.2017.07.006

Muschamp, J. W., Nemeth, C. L., Robison, A. J., Nestler, E. J., & Carlezon, W. A. (2012). ΔfosB enhances the rewarding effects of cocaine while reducing the pro-depressive effects of the kappa-opioid receptor agonist U50488. *Biological Psychiatry*, *71*(1), 44–50. https://doi.org/10.1016/j.biopsych.2011.08.011

Nagy, C., Vaillancourt, K., & Turecki, G. (2018). A role for activity-dependent epigenetics in the development and treatment of major depressive disorder. *Genes, Brain and Behavior*, *17*(3), e12446. https://doi.org/10.1111/gbb.12446

Nagy, Corina, Maitra, M., Tanti, A., Suderman, M., Théroux, J. F., Davoli, M. A., … Turecki, G. (2020). Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nature Neuroscience*, *23*(6), 771–781.

https://doi.org/10.1038/s41593-020-0621-y

Nakao, M., Matsui, S., Yamamoto, S., Okumura, K., Shirakawa, M., & Fujita, N. (2001). Regulation of transcription and chromatin by methyl-CpG binding protein MBD1. *Brain and Development*, *23*, S174–S176. https://doi.org/10.1016/S0387-7604(01)00348-5

Nelson, C. L., Milovanovic, M., Wetter, J. B., Ford, K. A., & Wolf, M. E. (2009). Behavioral sensitization to amphetamine is not accompanied by changes in glutamate receptor surface expression in the rat nucleus accumbens. *Journal of Neurochemistry*, *109*(1), 35–51. https://doi.org/10.1111/j.1471-4159.2009.05911.x

Nestler, E. J. (2008). Transcriptional mechanisms of addiction: Role of ΔFosB. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 363, pp. 3245–3255). Royal Society. https://doi.org/10.1098/rstb.2008.0067

Nestler, E. J. (2014a). Epigenetic mechanisms of drug addiction. *Neuropharmacology*, *76 Pt B*, 259–268. https://doi.org/10.1016/j.neuropharm.2013.04.004

Nestler, E. J. (2014b). Epigenetic mechanisms of drug addiction. *Neuropharmacology*, *76 Pt B*, 259–268. https://doi.org/10.1016/j.neuropharm.2013.04.004

Nielsen, D. A., Huang, W., Hamon, S. C., Maili, L., Witkin, B. M., Fox, R. G., … Moeller, F. G. (2012). Forced abstinence from cocaine self-administration is associated with DNA methylation changes in myelin genes in the corpus callosum: A preliminary study. *Frontiers in Psychiatry*, *3*(JUN), 1–9. https://doi.org/10.3389/fpsyt.2012.00060

Novikova, S. I., He, F., Bai, J., Cutrufello, N. J., Lidow, M. S., & Undieh, A. S. (2008). Maternal cocaine administration in mice alters DNA methylation and gene expression in hippocampal neurons of neonatal and prepubertal offspring. *PLoS ONE*, *3*(4). https://doi.org/10.1371/journal.pone.0001919

Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). <1999 DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development..pdf>, *99*, 247–257.

Otaka, M., Ishikawa, M., Lee, B. R., Liu, L., Neumann, P. A., Cui, R., … Dong, Y. (2013). Exposure to Cocaine Regulates Inhibitory Synaptic Transmission in the Nucleus Accumbens. *Journal of Neuroscience*, *33*(16), 6753–6758. https://doi.org/10.1523/JNEUROSCI.4577-12.2013

Patil, V., Ward, R. L., & Hesson, L. B. (2014). The evidence for functional non-CpG methylation in mammalian cells. *Epigenetics*, *9*(6), 823–828. https://doi.org/10.4161/epi.28741

Penrod, R. D., Anderson, E. M., & Cowan, C. W. (2018, August 1). Any Way You Splice It: New Molecular Mechanisms of Cocaine-Induced Alternative Gene Expression. *Biological Psychiatry*. Elsevier USA. https://doi.org/10.1016/j.biopsych.2018.06.004

Pol Bodetto, S., Carouge, D., Fonteneau, M., Dietrich, J. B., Zwiller, J., & Anglard, P. (2013). Cocaine represses protein phosphatase-1CB through DNA methylation and Methyl-CpG Binding Protein-2 recruitment in adult rat brain. *Neuropharmacology*, *73*, 31–40. https://doi.org/10.1016/j.neuropharm.2013.05.005

Pol Bodetto, S., Romieu, P., Sartori, M., Tesone-Coelho, C., Majchrzak, M., Barbelivien, A., … Anglard, P. (2014). Differential regulation of MeCP2 and PP1 in passive or voluntary administration of cocaine or food. *The International Journal of Neuropsychopharmacology / Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, *1*, 1–14.

https://doi.org/10.1017/S1461145714000972

Rehm, J., & Shield, K. D. (2019, January 1). Global Burden of Disease and the Impact of Mental and Addictive Disorders. *Current Psychiatry Reports*. Current Medicine Group LLC 1. https://doi.org/10.1007/s11920-019-0997-0

Renthal, W., Kumar, A., Xiao, G., Wilkinson, M., Covington, H., Maze, I., … Nestler, E. (2009). Genome-wide analysis of chromatin regulation by cocaine reveals a role for sirtuins. *Neuron*, *62*(3), 335–348. https://doi.org/10.1016/j.neuron.2009.03.026

Renthal, W., & Nestler, E. (2009). Histone acetylation in drug addiction. *Seminars in Cell & Developmental Biology*, *20*(4), 387–394. https://doi.org/10.1016/j.semcdb.2009.01.005

Robinson, T. E., & Kolb, B. (1999). Alterations in the morphology of dendrites and dendritic spines in the nucleus accumbens and prefrontal cortex following repeated treatment with amphetamine or cocaine. *The European Journal of Neuroscience*, *11*(5), 1598–1604.

Robison, A. J., & Nestler, E. J. (2011, November). Transcriptional and epigenetic mechanisms of addiction. *Nature Reviews Neuroscience*. Nat Rev Neurosci. https://doi.org/10.1038/nrn3111

Robledo, P., & Koob, G. F. (1993). Two discrete nucleus accumbens projection areas differentially mediate cocaine self-administration in the rat. *Behavioural Brain Research*, *55*(2), 159–166. https://doi.org/10.1016/0166-4328(93)90112-4

Romeo, K., Louault, Y., Cantaloube, S., Loiodice, I., Almouzni, G., & Quivy, J.-P. (2015). The SENP7 SUMO-Protease Presents a Module of Two HP1 Interaction Motifs that Locks HP1 Protein at Pericentric Heterochromatin. *Cell Reports*, *10*, 771–782. https://doi.org/10.1016/j.celrep.2015.01.004

Romieu, P., Host, L., Gobaille, S., Sandner, G., Aunis, D., & Zwiller, J. (2008). Histone deacetylase inhibitors decrease cocaine but not sucrose self-administration in rats. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(38), 9342–9348. https://doi.org/10.1523/JNEUROSCI.0379-08.2008

Rowlett, J. K., Platt, D. M., Yao, W. D., & Spealman, R. D. (2007). Modulation of heroin and cocaine self-administration by dopamine D1- and D2-like receptor agonists in rhesus monkeys. *Journal of Pharmacology and Experimental Therapeutics*, *321*(3), 1135–1143. https://doi.org/10.1124/jpet.107.120766

Sadri-Vakili, G., Kumaresan, V., Schmidt, H. D., Famous, K. R., Chawla, P., Vassoler, F. M., … Ghazaleh, S.-V. (2010). Cocaine-induced chromatin remodeling increases brain-derived neurotrophic factor transcription in the rat medial prefrontal cortex, which alters the reinforcing efficacy of cocaine. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(35), 11735–11744. https://doi.org/10.1523/JNEUROSCI.2328-10.2010

Saftig, P., & Lichtenthaler, S. F. (2015). The alpha secretase ADAM10: A metalloprotease with multiple functions in the brain. *Progress in Neurobiology*, *135*, 1–20. https://doi.org/10.1016/j.pneurobio.2015.10.003

Savell, K. E., Tuscher, J. J., Zipperly, M. E., Duke, C. G., Phillips, R. A., Bauman, A. J., … Day, J. J. (2020). A dopamine-induced gene expression signature regulates neuronal function and cocaine response. *Science Advances*, *6*(26). 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. *Annals of the New York Academy of Sciences*. Blackwell Publishing Inc. https://doi.org/10.1111/j.1749-6632.2009.05144.x

Schübeler, D. (2015). Function and information content of DNA methylation. *Nature*, *517*(7534), 321–326. https://doi.org/10.1038/nature14192

Schultz, W. (2007). Multiple Dopamine Functions at Different Time Courses. *Annual Review of Neuroscience*, *30*(1), 259–288. https://doi.org/10.1146/annurev.neuro.28.061604.135722

Skene, P. J., Illingworth, R. S., Webb, S., Kerr, A. R. W., James, K. D., Turner, D. J., … Bird, A. P. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Molecular Cell*, *37*(4), 457–468. https://doi.org/10.1016/j.molcel.2010.01.030

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., … Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, N.Y.)*, *324*(5929), 930–935. https://doi.org/10.1126/science.1170116

Tang, W.-X., Fasulo, W., Mash, D., & Hemby, S. (2003). Molecular profiling of midbrain dopamine regions in cocaine overdose victims. *Journal of Neurochemistry*, *85*(4), 911–924. https://doi.org/10.1046/j.1471-4159.2003.01740.x

Tannu, N. S., Howell, L. L., & Hemby, S. E. (2010). Integrative proteomic analysis of the nucleus accumbens in rhesus monkeys following cocaine self-administration. *Molecular Psychiatry*, *15*(2), 185–203. https://doi.org/10.1038/mp.2008.53

Tena, J. J., Alonso, M. E., de la Calle-Mustienes, E., Splinter, E., de Laat, W., Manzanares, M., & Gómez-Skarmeta, J. L. (2011). An evolutionarily conserved three-dimensional structure in the vertebrate Irx clusters facilitates enhancer sharing and coregulation. *Nature Communications*, *2*(1), 310. https://doi.org/10.1038/ncomms1301

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

Tian, W., Zhao, M., Li, M., Song, T., Zhang, M., Quan, L., … 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

Tiwari, V. K., McGarvey, K. M., Licchesi, J. D. F., Ohm, J. E., Herman, J. G., Schübeler, D., & Baylin, S. B. (2008). PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biology*, *6*(12), 2911–2927. https://doi.org/10.1371/journal.pbio.0060306

Turek-Plewa J. et al. (2005). The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cellular & Molecular Biology Letters*, *10*(4), 631–647.

Um, M., Whitt, Z., Revilla, R., Hunton, T., & Cyders, M. (2019). Shared Neural Correlates Underlying Addictive Disorders and Negative Urgency. *Brain Sciences*, *9*(2), 36. https://doi.org/10.3390/brainsci9020036

Uys, J., & LaLumiere, R. (2008). Glutamate: The New Frontier in Pharmacotherapy for Cocaine Addiction. *CNS & Neurological Disorders - Drug Targets*, *7*(5), 482–491. https://doi.org/10.2174/187152708786927868

Veremeyko, T., Siddiqui, S., Sotnikov, I., Yung, A., & Ponomarev, E. D. (2013). IL-4/IL-13-Dependent and Independent Expression of miR-124 and Its Contribution to M2 Phenotype of Monocytic Cells in Normal Conditions and during Allergic Inflammation. *PLoS ONE*, *8*(12), e81774. https://doi.org/10.1371/journal.pone.0081774

Volkow, N. D., Fowler, J. S., Wang, G. J., & Swanson, J. M. (2004, June). Dopamine in drug abuse and addiction: Results from imaging studies and treatment implications. *Molecular Psychiatry*. Mol Psychiatry. https://doi.org/10.1038/sj.mp.4001507

Volkow, N. D., Wang, G.-J., Telang, F., Fowler, J. S., Logan, J., Childress, A.-R., … Wong, C. (2006). Cocaine Cues and Dopamine in Dorsal Striatum: Mechanism of Craving in Cocaine Addiction. *Journal of Neuroscience*, *26*(24), 6583–6588. https://doi.org/10.1523/JNEUROSCI.1544-06.2006

Volkow, N. D., Wang, G. J., Fischman, M. W., Foltint, R. W., Fowler, J. S., Abumrad, N. N., … Shea, C. E. (1997). Relationship between subjective effects of cocaine and dopamine transporter occupancy. *Nature*, *386*(6627), 827–830. https://doi.org/10.1038/386827a0

Volkow, Nora D., Fowler, J. S., Wang, G. -J, Hitzemann, R., Logan, J., Schlyer, D. J., … Wolf, A. P. (1993). Decreased dopamine D2 receptor availability is associated with reduced frontal metabolism in cocaine abusers. *Synapse*, *14*(2), 169–177. https://doi.org/10.1002/syn.890140210

Volkow, Nora D., & Swanson, J. M. (2003, November). Variables that affect the clinical use and abuse of methylphenidate in the treatment of ADHD. *American Journal of Psychiatry*. Am J Psychiatry. https://doi.org/10.1176/appi.ajp.160.11.1909

Volkow, Nora D., Wang, G. J., Fowler, J. S., Tomasi, D., & Telang, F. (2011, September 13). Addiction: Beyond dopamine reward circuitry. *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.1010654108

Volkow, Nora D., Wang, G. J., Ma, Y., Fowler, J. S., Wong, C., Ding, Y. S., … Kalivas, P. (2005). Activation of orbital and medial prefrontal cortex by methylphenidate in cocaine-addicted subjects but not in controls: Relevance to addiction. *Journal of Neuroscience*, *25*(15), 3932–3939. https://doi.org/10.1523/JNEUROSCI.0433-05.2005

Volkow, Nora D., Wang, G. J., Telang, F., Fowler, J. S., Logan, J., Childress, A. R., … Wong, C. (2008). Dopamine increases in striatum do not elicit craving in cocaine abusers unless they are coupled with cocaine cues. *NeuroImage*, *39*(3), 1266–1273. https://doi.org/10.1016/j.neuroimage.2007.09.059

Walker, D. M., Cates, H. M., Loh, Y. H. E., Purushothaman, I., Ramakrishnan, A., Cahill, K. M., … Nestler, E. J. (2018). Cocaine Self-administration Alters Transcriptome-wide Responses in the Brain's Reward Circuitry. *Biological Psychiatry*, *84*(12), 867–880. https://doi.org/10.1016/j.biopsych.2018.04.009

Walton, E., Hass, J., Liu, J., Roffman, J. L., Bernardoni, F., Roessner, V., … Ehrlich, S. (2016). Correspondence of DNA Methylation Between Blood and Brain Tissue and Its Application to Schizophrenia Research. *Schizophrenia Bulletin*, *42*(2), 406–414. https://doi.org/10.1093/schbul/sbv074

Wang, J. Q., & McGinty, J. F. (1999). Glutamate-dopamine interactions mediate the effects of psychostimulant drugs. *Addiction Biology*. Addict Biol. https://doi.org/10.1080/13556219971641

Wang, K. H., Penmatsa, A., & Gouaux, E. (2015). Neurotransmitter and psychostimulant recognition by

the dopamine transporter. *Nature*, *521*(7552), 322–327. https://doi.org/10.1038/nature14431

Wang, L., Lv, Z., Hu, Z., Sheng, J., Hui, B., Sun, J., & Ma, L. (2010). Chronic cocaine-induced H3 acetylation and transcriptional activation of {CaMKIIalpha} in the nucleus accumbens is critical for motivation for drug reinforcement. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *35*(4), 913–928. https://doi.org/10.1038/npp.2009.193

Wang, X. Y., Yu, M., Ma, L., Wang, F. F., & Jiang, C. Y. (2019). Corticotrophin-releasing hormone neurons in the central amygdala mediate morphine withdrawal-induced negative emotions. *Sheng Li Xue Bao : [Acta Physiologica Sinica]*, *71*(6), 824–832. Retrieved from https://pubmed-ncbi-nlm-nih-gov.proxy3.library.mcgill.ca/31879737/

Wright, K. N., Hollis, F., Duclot, F., Dossat, A. M., Strong, C. E., Francis, T. C., … Kabbaj, M. (2015). Methyl Supplementation Attenuates Cocaine-Seeking Behaviors and Cocaine-Induced c-Fos Activation in a DNA Methylation-Dependent Manner. *Journal of Neuroscience*, *35*(23), 8948–8958. https://doi.org/10.1523/JNEUROSCI.5227-14.2015

Xenias, H. S., Ibáñez-Sandoval, O., Koós, T., Tepper, J. M., Tecuapetla, F., Koós, T., … Tepper, J. M. (2015). Are striatal tyrosine hydroxylase interneurons dopaminergic? *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *35*(16), 6584–6599. https://doi.org/10.1523/JNEUROSCI.0195-15.2015

Xue, G., Lu, Z., Levin, I. P., & Bechara, A. (2010). The impact of prior risk experiences on subsequent risky decision-making: The role of the insula. *NeuroImage*, *50*(2), 709–716. https://doi.org/10.1016/j.neuroimage.2009.12.097

Zachariou, V., Sgambato-Faure, V., Sasaki, T., Svenningsson, P., Berton, O., Fienberg, A. A., … Nestler, E. J. (2006). Phosphorylation of DARPP-32 at Threonine-34 is Required for Cocaine Action. *Neuropsychopharmacology*, *31*(3), 555–562. https://doi.org/10.1038/sj.npp.1300832

Zhang, X., Lee, T. H., Davidson, C., Lazarus, C., Wetsel, W. C., & Ellinwood, E. H. (2007). Reversal of cocaine-induced behavioral sensitization and associated phosphorylation of the NR2B and GluR1 subunits of the NMDA and AMPA receptors. *Neuropsychopharmacology*, *32*(2), 377–387. https://doi.org/10.1038/sj.npp.1301101

Zhao, Q., Hou, J., Chen, B., Shao, X., Zhu, R., Bu, Q., … Cen, X. (2015). Prenatal cocaine exposure impairs cognitive function of progeny via insulin growth factor II epigenetic regulation. *Neurobiology of Disease*, *82*, 54–65. https://doi.org/10.1016/j.nbd.2015.05.014

Zhou, Zhaolan, Hong, E. J., Cohen, S., Zhao, W.-N., Ho, H.-Y. H., Schmidt, L., … Greenberg, M. E. (2006). Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron*, *52*(2), 255–269. https://doi.org/10.1016/j.neuron.2006.09.037

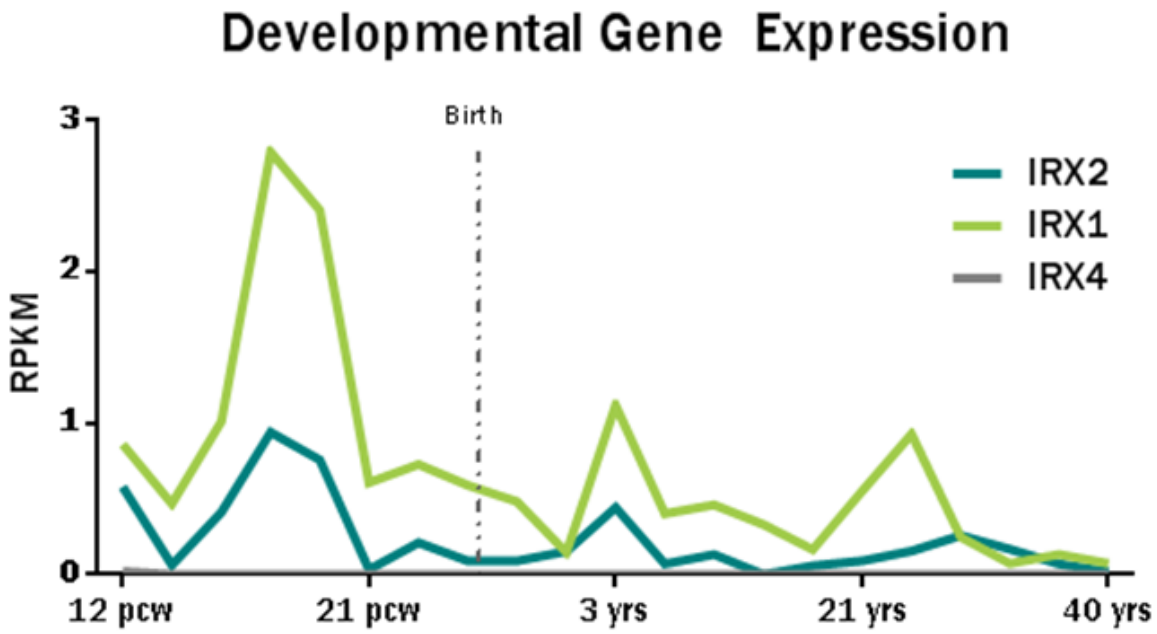
Zhou, Zhifeng, Yuan, Q., Mash, D. C., & Goldman, D. (2011). Substance-specific and shared transcription and epigenetic changes in the human hippocampus chronically exposed to cocaine and alcohol. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(16), 6626–6631.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

Zoghbi, H. Y., Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., & Francke, U. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, *23*(2), 185–188. https://doi.org/10.1038/13810

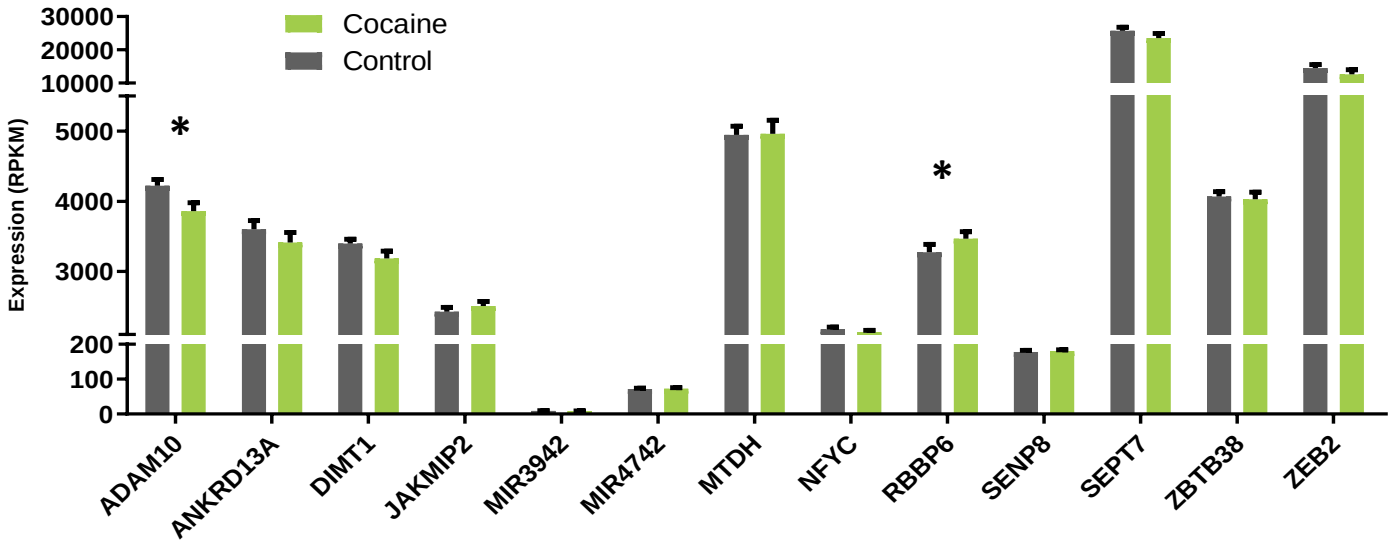
Zorrilla, E. P., & Koob, G. F. (2019, May 22). Impulsivity derived from the dark side: Neurocircuits that contribute to negative urgency. *Frontiers in Behavioral Neuroscience*. Frontiers Media S.A. https://doi.org/10.3389/fnbeh.2019.00136

APPENDICES

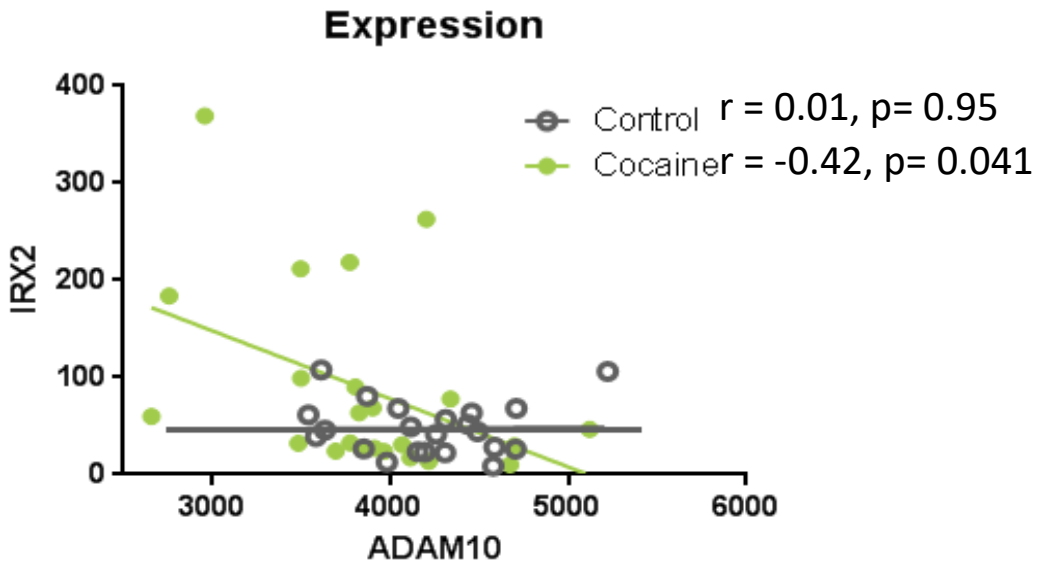
APPENDIX 1. SUPPLEMENTARY FIGURES



Supplemental Figure 1. Expression of the members of the IRXA gene cluster throughout development. Data accessed from the BrainSpan Human Brain Atlas.



Supplemental Figure 2. Differential Expression of Predicted IRX2 Target Genes in Caudate. We used MotifMap (http://motifmap.ics.uci.edu/) to identify potential regulatory targets of the IRX2 transcription factor in humans, based on in silico predictions. 13 potential targets were detected within the RNA sequencing dataset from the discovery cohort, and two genes were nominally differentially expressed between the cocaine group and controls (ADAM10, RBBP6). Data represent mean ± s.e.m.
* Two tailed Student's t-test $p < 0.05$



Supplemental Figure 3. IRX2 expression is negatively associated with ADAM10 expression in the caudate of cocaine-dependent individuals, but not controls. Data analysed with Pearson's correlations.

APPENDIX 2. SIGNIFICANT CONTRIBUTIONS OF THESIS AUTHOR TO OTHER PROJECTS

- Nagy C*, **Vaillancourt K*,** Turecki G. (2018) *A role for activity dependent epigenetics in the development and treatment of major depressive disorder*. Genes, Brain and Behaviour. 17(3):e12446

- Lutz PE, Tanti A, Gasecka A, Barnett-Burns S, Kim JJ, Zhou Y, Chen GG, Wakid M, Shaw M, Almeida D, Chay MA, Yang J, Larivière V, M'Boutchou MN, van Kempen LC, Yerko V, Prud'homme J, Davoli MA, **Vaillancourt K**, Théroux JF, Bramoullé A, Zhang TY, Meaney MJ, Ernst C, Côté D, Mechawar N, Turecki G. (2017) *Association of a History of Child Abuse with Impaired Myelination in the Anterior Cingulate Cortex: Convergent Epigenetic, Transcriptional, and Morphological Evidence*. American Journal of Psychiatry. 174(12): 1185-1194

- Chen GG, Gross JA, Lutz PE, **Vaillancourt K**, Maussion G, Bramoulle A, Théroux JF, Gardini ES, Ehlert U, Bourret G, Masurel A, Lepage P, Mechawar N, Turecki G, Ernst C. (2017) *Medium throughput bisulfite sequencing for accurate detection of 5-methylcytosine and 5-hydroxymethycytosine.* BMC Genomics. 18(1):96

- Nagy C, Maheu M, Lopez JP, **Vaillancourt K**, Cruceanu C, Gross JA, Arnovitz M, Mechawar N, Turecki G. (2015) *Effects of postmortem interval on biomolecule integrity in the brain.* Journal of Neuropathology and Experimental Neurology. 74(5):459-69

- Elizabeth S. Chen, Carolina O. Gigek, Jill A. Rosenfeld, Alpha B. Diallo, Gilles Maussion, Gang Chen, **Kathryn Vaillancourt**, Liam Crapper, Lisa G. Shaffer, Carl Ernst. (2014) *Molecular convergence of neurodevelopmental disorders*. American Journal of Human Genetics. 95(9): 490-508

- Chen G, Diallo A, Poujol R, Nagy C, Staffa A, **Vaillancourt K**, Lutz PE, Ota VK, Mash DCm Turecki G, Ernst C. (2014). *Bis-QC: An operational pipeline for multiplexed bisulfite sequencing*. BMC Genomics. 15(290)

PERMISSIONS TO REPRINT

1. MDPI

No special permission is required to reuse all or part of article published by MDPI, including figures and tables. For articles published under an open access Creative Common CC BY license, any part of the article may be reused without permission provided that the original article is clearly cited. Reuse of an article does not imply endorsement by the authors or MDPI.

(www.mdpi.com/openaccess)

2. NATURE PUBLISHING GROUP

Ownership of copyright in the article remains with the Authors, and provided that, when reproducing the Contribution or extracts from it, the Authors acknowledge first and reference publication in the Journal, the Authors retain the following non-exclusive rights: a) To reproduce the Contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

b) They and any academic institution where they work at the time may reproduce the Contribution for the purpose of course teaching.

c) To post a copy of the Contribution as accepted for publication after peer review (in Word or Tex format) on the Authors' own web site or institutional repository, or the Authors' funding body's designated archive, six months after publication of the printed or online edition of the Journal, provided that they also give a hyperlink from the Contribution to the Journals web site.

d) To reuse figures or tables created by them and contained in the Contribution in other works created by them.

The above use of the term 'Contribution' refers to the author's own version, not the final version as published in the Journal.

(https://www.nature.com/nature-research/reprints-and-permissions/permissions-requests)